

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 1, 2005, 17:24:23 ; Search time 4399 Seconds  
(without alignments)  
2478.387 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMRFKVRMEG.....EDYTIVEQYTERGHHFL 225

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO spool/US1006922/runat\_30062005\_110616\_19088/app query.fasta\_1.391  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US1006922 @CGN 1 1 3731 @runat\_30062005\_110616\_19088 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1214	100.0	678	6 AR183915	AR183915 Sequence
2	1214	100.0	678	6 AX172854	AX172854 Sequence
3	1214	100.0	678	6 AX207715	AX207715 Sequence
4	1214	100.0	678	6 AX233581	AX233581 Sequence

5	1214	100.0	678	6 AX370404	AX370404 Sequence
6	1214	100.0	678	6 AX824725	AX824725 Sequence
7	1214	100.0	859	3 AF168419	AF168419 Discosoma
8	1214	100.0	859	6 AX463698	AX463698 Sequence
9	1210	99.7	681	6 AX233584	AX233584 Sequence
10	1210	99.7	713	6 AX233627	AX233627 Sequence
11	1210	99.7	723	6 AR527331	AR527331 Sequence
12	1210	99.7	2721	6 CQ882115	CQ882115 Sequence
13	1210	99.7	2772	6 CQ882117	CQ882117 Sequence
14	1210	99.7	4692	6 AX463702	AX463702 Sequence
15	1210	99.7	6893	6 AX823860	AX823860 Sequence
16	1210	99.7	8811	12 AY569779	AY569779 Cloning v
17	1210	99.7	9320	6 AX663075	AX663075 Sequence
18	1210	99.7	12404	12 AY569780	AY569780 Cloning v
19	1207	99.4	678	6 AX370406	AX370406 Sequence
20	1196	98.5	666	6 AX348043	AX348043 Sequence
21	1196	98.5	666	6 AX353910	AX353910 Sequence
22	1196	98.5	711	3 AY679107	AY679107 Discosoma
23	1193	98.3	678	3 AF545828	AF545828 Discosoma
24	1191	98.1	678	6 AX370408	AX370408 Sequence
25	1191	98.1	678	6 AX824732	AX824732 Sequence
26	1191	98.1	921	3 AY679106	AY679106 Discosoma
27	1186	97.7	675	6 AX824731	AX824731 Sequence
28	1186	97.7	1050	6 AX666133	AX666133 Sequence
29	1186	97.7	4488	6 CQ849509	CQ849509 Sequence
30	1186	97.7	6423	12 AY613997	AY613997 Cloning v
31	1186	97.7	7616	6 CQ849511	CQ849511 Sequence
32	1184	97.5	898	6 AX686888	AX686888 Sequence
33	1165	96.0	10141	12 AY342347	AY342347 Red H-Pel
34	1165	96.0	10276	12 AY342348	AY342348 Red H-Sti
35	1165	96.0	10481	12 AY490568	AY490568 UAS-Red S
36	1121	92.3	681	12 AF506025	AF506025 Synthetic
37	1121	92.3	1395	12 AF506026	AF506026 Synthetic
38	1119	92.2	678	6 AX824729	AX824729 Sequence
39	1085.5	89.4	876	3 AF272711	AF272711 Discosoma
40	1085.5	89.4	876	6 AX686894	AX686894 Sequence
41	1021	84.1	678	12 AF506027	AF506027 Synthetic
42	1021	84.1	696	12 AB166761	AB166761 Synthetic
43	805	66.3	669	6 AX699925	AX699925 Sequence
44	804	66.2	666	3 AY646074	AY646074 Acropora
45	800	65.9	663	6 AX699859	AX699859 Sequence

ALIGNMENTS

RESULT 1  
AR183915  
LOCUS AR183915 Sequence 678 bp DNA linear PAT 20-APR-2002  
DEFINITION Sequence 6 from patent US 6342379.  
ACCESSION AR183915  
VERSION AR183915.1 GI:20227884  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 678)  
AUTHORS Tsien,R.Y. and Gonzalez,J.E. III.  
TITLE Detection of transmembrane potentials by optical methods  
JOURNAL Patent: US 6342379-A 6 29-JAN-2002;  
FEATURES Location/Qualifiers  
source  
1. .678  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Alignment Scores:  
Pred. No.: 2.62e-119 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

Not a  
Nucleotide  
1022e





FEATURES  
source Aurora Biosciences Corporation (US)  
Location/Qualifiers  
1. .678  
/organism="Discosoma sp."  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:86600"

ORIGIN

Alignment Scores:  
Pred. No.: 2.62e-119 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x AX207715 (1-678)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
|||||  
Db 1 ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGGA 60

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
|||||  
Db 61 ACGTCAATGGGCACGAGTTTGAATAGAGCGCAAGGAGGAGGGGAGGCCATACGAAGGC 120

QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
|||||  
Db 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180

QY 61 LeuSerProGlnPheGlnTyrGlySerLysValThrLysHisProAlaAspIlePro 80  
|||||  
Db 181 TTGTCACCACAAATTTTCAGTATGGAAGCAAGGTATATGTCGAAGCACCCCTGCCGACATACCA 240

QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
|||||  
Db 241 GACTATAAAAAGCTGTCTTCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300

QY 101 AspGlyGlyValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
|||||  
Db 301 GACGGTGGCGTCTACTGTAAACCCAGGATTTCCAGTTTGCAGGATGGCTGTTTCATCTAC 360

QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
|||||  
Db 361 AAGTCAAGTTTCATTTGGCGTGAACTTTCCCTCCGATGGACCTGTTATGCAAAAGAAGACA 420

QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
|||||  
Db 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480

QY 161 IleHisLysAlaLeuLysLeuLysAspGlyHisTyrLeuValGluPheLysSerIle 180  
|||||  
Db 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTTCATACCTAGTTGAATTCAAAAGTATT 540

QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrValAspSerLysLeuAsp 200  
|||||  
Db 541 TACATGGCAAGAAGCCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
|||||  
Db 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGACGAGTATGAAAGAACCGAGGGACGC 660

QY 221 HisHisLeuPheLeu 225  
|||||

RESULT 4  
AX233581  
LOCUS AX233581 678 bp DNA linear PAT 11-SEP-2001  
DEFINITION Sequence 5 from Patent WO0162919.  
ACCESSION AX233581  
VERSION AX233581.1 GI:15593305  
KEYWORDS

SOURCE Discosoma sp.  
ORGANISM Discosoma sp.  
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;  
Corallimorpharia; Discosomatidae; Discosoma.

REFERENCE 1  
AUTHORS Nelson,D., Zamaira,E. and Tsien,R.  
TITLE Modified fluorescent proteins  
JOURNAL Patent: WO 0162919-A 5 30-AUG-2001;  
Aurora Biosciences Corporation (US)

FEATURES  
source Location/Qualifiers  
1. .678  
/organism="Discosoma sp."  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:86600"  
/note="red"

ORIGIN

Alignment Scores:  
Pred. No.: 2.62e-119 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x AX233581 (1-678)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
|||||  
Db 1 ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGGA 60

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
|||||  
Db 61 ACGTCAATGGGCACGAGTTTGAATAGAGCGCAAGGAGGAGGCCATACGAAGGC 120

QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
|||||  
Db 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180

QY 61 LeuSerProGlnPheGlnTyrGlySerLysValThrLysHisProAlaAspIlePro 80  
|||||  
Db 181 TTGTCACCACAAATTTTCAGTATGGAAGCAAGGTATATGTCGAAGCACCCCTGCCGACATACCA 240

QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
|||||  
Db 241 GACTATAAAAAGCTGTCTTCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300

QY 101 AspGlyGlyValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
|||||  
Db 301 GACGGTGGCGTCTACTGTAAACCCAGGATTTCCAGTTTGCAGGATGGCTGTTTCATCTAC 360

QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
|||||  
Db 361 AAGTCAAGTTTCATTTGGCGTGAACTTTCCCTCCGATGGACCTGTTATGCAAAAGAAGACA 420

QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
|||||  
Db 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480

QY 161 IleHisLysAlaLeuLysLeuLysAspGlyHisTyrLeuValGluPheLysSerIle 180  
|||||  
Db 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTTCATACCTAGTTGAATTCAAAAGTATT 540

QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrValAspSerLysLeuAsp 200  
|||||  
Db 541 TACATGGCAAGAAGCCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
|||||  
Db 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGACGAGTATGAAAGAACCGAGGGACGC 660

QY 221 HisHisLeuPheLeu 225  
|||||



```
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
|
|
|
Db 481 ATCCACAAGGCCCTGAAGCTGAAGGACGGCGCCACTACTCTGGTGAGTTCAAGTCCATC 540

QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
|
|
|
Db 541 TACATGGCCAAAGAGCCCGTGCAGCTGCCCGCTACTACTACGTGGACTCCAAAGCTGGAC 600

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGlnTyrGluArgThrGluGlyArg 220
|
|
|
Db 601 ATCACCTCCCAACAGGAGACTACACCATCGTGGAGCAGTACGAGCGCACCGAGGGCCGC 660

QY 221 HisHisLeuPheLeu 225
|
|
|
Db 661 CACCACCTGTTTCCTG 675

RESULT 7
AF168419 859 bp mRNA linear INV 27-JUL-2001
LOCUS Discosoma sp. fluorescent protein FP583 mRNA, complete cds.
ACCESSION AF168419
VERSION AF168419.2 GI:7105733
KEYWORDS
SOURCE Discosoma sp.
ORGANISM Discosoma sp.
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;
Corallimorpharia; Discosomatidae; Discosoma.
REFERENCE
AUTHORS Matz,M.V., Fradkov,A.F., Labas,Y.A., Savitsky,A.P., Zارايسكى,A.G.,
Markelov,M.L. and Lukyanov,S.A.
TITLE Fluorescent proteins from nonbioluminescent Anthozoa species
JOURNAL Nat. Biotechnol. 17 (10), 969-973 (1999)
MEDLINE 99436614
PUBMED 10504696
REFERENCE
AUTHORS Matz,M.V., Fradkov,A.F., Labas,Y.A., Savitsky,A.P., Zارايسكى,A.G.,
Markelov,M.L. and Lukyanov,S.A.
TITLE Direct Submission
JOURNAL Submitted (14-JUL-1999) Institute of Bioorganic Chemistry RAS,
Miklukho-Maklaya 16/10, Moscow 117871, Russia
REFERENCE
AUTHORS Matz,M.V., Fradkov,A.F., Labas,Y.A., Savitsky,A.P., Zارايسكى,A.G.,
Markelov,M.L. and Lukyanov,S.A.
TITLE Direct Submission
JOURNAL Submitted (25-FEB-2000) Institute of Bioorganic Chemistry RAS,
Miklukho-Maklaya 16/10, Moscow 117871, Russia
REMARK Sequence update by submitter
COMMENT On Feb 25, 2000 this sequence version replaced gi:6090866.
FEATURES
source
1. .859
/organism="Discosoma sp."
/mol_type="mRNA"
/db_xref="taxon:86600"
. 54..731
/note="GFP-like protein; orange-red"
/codon_start=1
/product="fluorescent protein FP583"
/protein_id="AAF03369.1"
/db_xref="GI:6090867"
/translation="MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGERPYEGHNTV
KLKVTKGGLPFAWDILSPQFYGSKVYVKHPADIPDYKLSFEGFKWERVMNFEDG
GVVTVTQDSSLQDCFIYKVKRFIGVNFPSDGPVMQKTMGWEASTERLYPRDGVLKGE
IHKALKLKDGGHYLVEFKSIYMAKKPVQLPGYVVVDSKLDITSHNEDYTIIVEQYERTE
GRHHLFL"

ORIGIN
Alignment Scores:
Pred. No.: 3.53e-119 Length: 859
Score: 1214.00 Matches: 225
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
```

```
DB: 3 Gaps: 0
US-10-006-922A-12 (1-225) x AF168419 (1-859)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
|
|
|
Db 54 ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTTCATGAGGTTTTAAGTTTCGCATGGAAGGA 113

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40
|
|
|
Db 114 ACGGTCAATGGGCACGAGTTTGAATAGAAGCGGAAGGAGAGGGAGGCCCATACGAAGGC 173

QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60
|
|
|
Db 174 CACAATACCGTAAAGCTTAAGGTAAACCAAGGGGGGACCTTTGCCATTGCTTGGGATATT 233

QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80
|
|
|
Db 234 TTGTCAACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 293

QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100
|
|
|
Db 294 GACTATAAAAAGCTGTCTATTTCTCTGAAGGATTTAATGGGAAAGSGTCATGAACCTTTGAA 353

QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
|
|
|
Db 354 GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 413

QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
|
|
|
Db 414 AAGGTCAAGTTCATTGGCGTGAACCTTCCTTCCGATGGACCTGTTATGCAAAAAGAAGACA 473

QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
|
|
|
Db 474 ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCCTCGTATGGCGTGTGAAAGGAGAG 533

QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
|
|
|
Db 534 ATTCTAAGGCTCTGAAGCTGAAGACGGTGGTGCATTACCTAGTTGAATTCAAAAGTATT 593

QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
|
|
|
Db 594 TACATGGCAAGAAGCCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 653

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGlnTyrGluArgThrGluGlyArg 220
|
|
|
Db 654 ATAACAAGCCACAAACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCAGGAGGRCG 713

QY 221 HisHisLeuPheLeu 225
|
|
|
Db 714 CACCATCTGTTTCCTT 728

RESULT 8
AX463698 859 bp DNA linear PAT 15-JUL-2002
LOCUS Sequence 12 from Patent WO0248338.
ACCESSION AX463698
VERSION AX463698.1 GI:21886457
KEYWORDS
SOURCE Discosoma sp.
ORGANISM Discosoma sp.
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;
Corallimorpharia; Discosomatidae; Discosoma.
REFERENCE
AUTHORS Lichtenberg-Frat,H.
TITLE Yeast strain for testing the geno- and cytotoxicity of complex
environmental contamination
JOURNAL Patent: WO 0248338-A 12 20-JUN-2002;
Lichtenberg-Frat, Hella (DE)
FEATURES
source
1. .859
/organism="Discosoma sp."
/mol_type="unassigned DNA"
```



CDS /db\_xref="taxon:86600" 54..731 /note="unnamed protein product" /codon\_start=1 /protein\_id="CAD42147.1" /db\_xref="GI:21886458" /translation="MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTV KLKVTGGPLPFAWDILSPQFYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDG GVVTVTQDSSLQDGCIFYKVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGE IHKALKLDGGHYLVEFKSIYMAKKPVQLPGYYVDSKLDITSHNEDYTIQEYRTE GRHHLFL"

ORIGIN Alignment Scores: Pred. No.: 3.53e-119 Length: 859 Score: 1214.00 Matches: 225 Percent Similarity: 100.00% Conservative: 0 Best Local Similarity: 100.00% Mismatches: 0 Query Match: 100.00% Indels: 0 DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x AX463698 (1-859)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 54 ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTCCGATGGAAGGA 113

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 114 ACGGTCAATGGGCACGAGTTTGAATAGAACGGCGAAGGAGAGGGGAGGCCATACGAAGGC 173

QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 174 CACAATACCGTAAAGCTTAAGTTAACCAAGGGGGACCTTTGCCATTTGCTGGGATATT 233

QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 234 TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA 293

QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
Db 294 GACTATAAAAAGCTGTCAATTTCCTGAAGGATTTAAATGGGAAAGGTCATGAACTTTGAA 353

QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 354 GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGATTCCAGTTTCAGGATGGCTGTTTCATCTAC 413

QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 414 AAGGTCAAGTTTCATTGGCGTGAACCTTCCTTCCGATGGACCTGTTATGCAAAAGAAAGACA 473

QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 474 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 533

QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 534 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTGCATTACCTAGTTGAATTCAAAAGTATT 593

QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 594 TACATGGCAAGAAGCCCTGTCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 653

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 654 ATAACAAGCCACACGAGACTATACAATCGTTGAGCAGTATGAAAGAACCCGAGGGACGC 713

QY 221 HisHisLeuPheLeu 225  
Db 714 CACCATCTGTTCCCTT 728

RESULT 9  
AX2333584

LOCUS AX2333584 681 bp DNA linear PAT 11-SEP-2001  
DEFINITION Sequence 8 from Patent WO0162919.  
ACCESSION AX2333584  
VERSION AX2333584.1 GI:15593307  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE 1 Nelson,D., Zamaira,E. and Tsien,R.  
AUTHORS Modified fluorescent proteins  
TITLE Patent: WO 0162919-A 8 30-AUG-2001;  
JOURNAL Aurora Biosciences Corporation (US)  
FEATURES Location/Qualifiers  
source 1..681  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Mutant Anthozoan red fluorescent protein"  
CDS 1..681  
/note="unnamed protein product"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="CAC69734.1"  
/db\_xref="GI:15593308"  
/translation="MVRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNT VKLKVTKGGPLPFAWDILSPQFYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFED GVVTVTQDSSLQDGCIFYKVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGE EIHKALKLDGGHYLVEFKSIYMAKKPVQLPGYYVDSKLDITSHNEDYTIQEYERT EGRHHLFL"

ORIGIN Alignment Scores: Pred. No.: 6.99e-119 Length: 681 Score: 1210.00 Matches: 224 Percent Similarity: 100.00% Conservative: 1 Best Local Similarity: 99.56% Mismatches: 0 Query Match: 99.67% Indels: 0 DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x AX2333584 (1-681)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 4 GTGAGGAGCAGCAAGAACGTGATCAAGGAGTTTCATGAGTTCAAGTTCAGGTCGATGGAGGC 63

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 64 ACCGTGAACGGCCACGAGTTTCGAGATCGAGGGCGAGGGCGAGGCCCTACGAGGC 123

QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 124 CACAACACCGTGAAGCTTAAGGTGACCAAGGGCGGCCCTGCCCTCGCCTGGGACATC 183

QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 184 CTGAGCCCCCAGTTCAGTACGGCAGCAAGGTGTACGTGAAGCACCCCGCCGACATCCCC 243

QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
Db 244 GACTACAAGAAGCTGAGCTTCCCCGAGGCTTCAAGTGGGAGAGGTGATGAACCTTCGAG 303

QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 304 GACGGCGCGTGGTGACCGGTGACCCAGCAGCAGCCCTGCAGGACGGCTGCTTCATCTAC 363

QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 364 AAGGTGAAGTTCATCGGCGTGAACITCCCCAGCGACGGCCCCCGTGTATGCAGAAAGACC 423

QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 424 ATGGGCTGGGAGGCGCTCCACCGAGCGCCTGTACCCCCCGACGGCGCTGCTGAAGGGCGAG 483

QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 484 ATCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACCTGGTGGAGTTCAAGTCCATC 543

QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 544 TACATGGCCCAAGAACGCCGTGCAGCTGCCCGGTACTACTACGTGGACTCCAAGCTGGAC 603

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 604 ATCACCAGCCACAAACGAGGACTACACCATCGTGGAGCAGTACGAGAGGACCGAGGGCAGG 663

QY 221 HisHisLeuPheLeu 225  
Db 664 CACCACCTGTTCTCTG 678

RESULT 10  
AX2333627  
LOCUS AX2333627 713 bp DNA linear PAT 11-SEP-2001  
DEFINITION Sequence 51 from Patent WO0162919.  
ACCESSION AX2333627  
VERSION AX2333627.1 GI:15593330  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
other sequences; artificial sequences.

REFERENCE 1  
AUTHORS Nelson,D., Zamaira,E. and Tsien,R.  
TITLE Modified fluorescent proteins  
JOURNAL Patent: WO 0162919-A 51 30-AUG-2001;  
Aurora Biosciences Corporation (US)

FEATURES  
source  
1..713  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Mutant Anthozoan red fluorescent protein"

ORIGIN  
Alignment Scores:  
Pred. No.: 7.41e-119 Length: 713  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x AX2333627 (1-713)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 22 GTGAGGAGCAGCAAGACGTGATCAAGGAGTTTCATGAGGTTCAAGGTGCGCATGGAGGC 81

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 82 ACCGTGAACGCCACGAGTTTCGAGATCGAGGGCGAGGGCGAGGGCCCTACGAGGC 141

QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 142 CACAACACCGTGAAGCTTAAGTGACCAAGGGCGGCCCTTCGCTTGGGACATC 201

QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 202 CTGAGCCCCCAGTTCCAGTACGGCAGCAAGGTGACGTGAAGCACCCCGCGACATCCCC 261

QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
Db 262 GACTACAAGAAGCTGAGCTTCCCCGAGGGCTTCAAGTGGGAGAGGTGATGAACCTTCGAG 321

QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 322 GACGGCGCGTGGTGACCGTGACCCAGGACAGCAGCCTGCAGGACGGGTGCTTTCATCTAC 381

QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 382 AAGGTGAAGTTTCATCGCGTGAACTTCCCCAGCGACGGCCCCGTGATGCAGAAGAAGACC 441

QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 442 ATGGGCTGGGAGGCCCTCCACCGAGCGCTGTACCCCCGCGACGGCGTGTGAAGGGCGAG 501

QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 502 ATCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACCTGGTGGAGTTCAAGTCCATC 561

QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 562 TACATGGCCCAAGAACGCCGTGCAGCTGCCCGGTACTACTACGTGGACTCCAAGCTGGAC 621

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 622 ATCACCAGCCACAAACGAGGACTACACCATCGTGGAGCAGTACGAGAGGACCGAGGGCAGG 681

QY 221 HisHisLeuPheLeu 225  
Db 682 CACCACCTGTTCTCTG 696

RESULT 11  
AR527331  
LOCUS AR527331 723 bp DNA linear PAT 08-OCT-2004  
DEFINITION Sequence 1 from patent US 6723537.  
ACCESSION AR527331  
VERSION AR527331.1 GI:53914309  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 723)  
AUTHORS Peelle,B.  
TITLE Directed evolution of protein in mammalian cells  
JOURNAL Patent: US 6723537-A 1 20-APR-2004;  
FEATURES  
source  
1..723  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Alignment Scores:  
Pred. No.: 7.54e-119 Length: 723  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x AR527331 (1-723)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 4 GTGCGCTCCTCCAAGAACGTTCATCAAGGAGTTTCATGCGCTTCAAGGTGCGCATGGAGGC 63

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 64 ACCGTGAACGCCACGAGTTTCGAGATCGAGGGCGAGGGCGAGGGCCCTACGAGGC 123

QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 124 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTCGCTTGGGACATC 183

QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 184 CTGTCCCCCAGTTCCAGTACGGCTCCAAGGTGTACGTGAAGCACCCCGCGACATCCCC 243

QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100

*Handwritten signature*

Db	244	GACTACAAGAGCTGTCCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACCTTCGAG	303
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	304	GACGGCGGCGTGGTGACCGTGACCCAGGACTCCTCCCTGCAGGACGGCTGCTTCATCTAC	363
QY	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
Db	364	AAGGTGAAGTTTCATCGCGGTGAACCTTCCCTCCGACGGCCCCGTAATGCAGAAGACC	423
QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	424	ATGGGCTGGGAGGCTCCACCGAGCGCTGTACCCCGCGACGGCGTGCTGAAGGGCGAG	483
QY	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db	484	ATCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACCTGGTGGAGTTCAAGAGTATC	543
QY	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	544	TACATGGCCAAAGACCGGTGCAGCTGCCCGGTACTACTACGTGGACTCCAAGCTGGAC	603
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	604	ATCACCTCCCAACAGGAGCTACACCATCGTGGAGCAGTACGAGCGCACCGAGGGCCGC	663
QY	221	HisHisLeuPheLeu 225	
Db	664	CACCACCTGTTTCCTG 678	
RESULT 12			
LOCUS	CQ882115	Sequence 1 from Patent WO2004083445.	2721 bp DNA linear PAT 11-OCT-2004
DEFINITION	CQ882115		
ACCESSION	CQ882115.1	GI:54034825	
VERSION			
KEYWORDS			
SOURCE		synthetic construct	
ORGANISM		synthetic construct	
REFERENCE	1	other sequences; artificial sequences.	
AUTHORS		Chavancy,G., Couble,P., Durand,B., Grenier,A.M., Horard,B., Julien,E., Mauchamp,B., Nony,P., Prudhomme,J.C. and Royer,C.	
TITLE		Nucleic acid controlling the expression of a useful polypeptide in the posterior silk glands of a lepidoptera and application thereof	
JOURNAL		Patent: WO 2004083445-A 1 30-SEP-2004;	
		Centre National De La Recherche Scientifique-CNRS (FR); UNIVERSITE CLAUDE BERNARD - LYON 1 (FR); Institut National de la Recherche Agronomique (INRA) (FR)	
FEATURES		Location/Qualifiers	
source	1..2721	/organism="synthetic construct"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:32630"	
		/note="Description de la s quence artificielle : s quence de fusion"	
ORIGIN			
Alignment Scores:			
Pred. No.:	4.01e-118	Length:	2721
Score:	1210.00	Matches:	224
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.56%	Mismatches:	0
Query Match:	99.67%	Indels:	0
DB:	6	Gaps:	0
US-10-006-922A-12 (1-225) x CQ882115 (1-2721)			
QY	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
Db	2044	GTGCGCTCCTCCAAGAACGTTCATCAAGGAGTTTCATCGCTTCAAGGTGCGCATGGAGGGC	2103
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40

Db	2104	ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCGCCCTACGAGGGC	2163
QY	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle	60
Db	2164	CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCCCTTGGCCTTGGGACATC	2223
QY	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	2224	CTGTCCCCCAGTTCCAGTACGGCTCCAAGGTGTACGTGAAGCACCCCGCGACATCCCC	2283
QY	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
Db	2284	GACTACAAGAAAGCTGTCTTCCCGAGGGCTTCAAGTGGGAGCGGTGATGAACCTTCGAG	2343
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	2344	GACGGCGGCGTGGTGACCGTGACCCAGGACTCCTCCCTGCAGGACGGCTGCTTCATCTAC	2403
QY	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
Db	2404	AAGGTGAAGTTTCATCGGCGTGAACTTCCCTCCGACGGCCCCGTAATGCAGAAGACC	2463
QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	2464	ATGGGCTGGGAGGCCTCCACCGAGCGCTGTACCCCGCGACGGCGTGCTGAAGGGCGAG	2523
QY	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db	2524	ATCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACCTGGTGGAGTTCAAGTCCATC	2583
QY	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	2584	TACATGGCCAAAGACCGGTGCAGCTGCCCGGTACTACTACTGGAGCTCCAAGCTGGAC	2643
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	2644	ATCACCTCCCAACAGGAGCTACACCATCGTGGAGCAGTACGAGCGCACCGAGGGCCGC	2703
QY	221	HisHisLeuPheLeu 225	
Db	2704	CACCACCTGTTTCCTG 2718	
RESULT 13			
LOCUS	CQ882117	Sequence 3 from Patent WO2004083445.	2772 bp DNA linear PAT 11-OCT-2004
DEFINITION	CQ882117		
ACCESSION	CQ882117.1	GI:54034827	
VERSION			
KEYWORDS		synthetic construct	
SOURCE		synthetic construct	
ORGANISM		other sequences; artificial sequences.	
REFERENCE	1	Chavancy,G., Couble,P., Durand,B., Grenier,A.M., Horard,B., Julien,E., Mauchamp,B., Nony,P., Prudhomme,J.C. and Royer,C.	
AUTHORS		Nucleic acid controlling the expression of a useful polypeptide in the posterior silk glands of a lepidoptera and application thereof	
TITLE		Patent: WO 2004083445-A 3 30-SEP-2004;	
JOURNAL		Centre National De La Recherche Scientifique-CNRS (FR); UNIVERSITE CLAUDE BERNARD - LYON 1 (FR); Institut National de la Recherche Agronomique (INRA) (FR)	
FEATURES		Location/Qualifiers	
source	1..2772	/organism="synthetic construct"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:32630"	
		/note="Description de la s quence artificielle : s quence de fusion"	
ORIGIN			
Alignment Scores:			
Pred. No.:	4.1e-118	Length:	2772
Score:	1210.00	Matches:	224
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.56%	Mismatches:	0
Query Match:	99.67%	Indels:	0
DB:	6	Gaps:	0
US-10-006-922A-12 (1-225) x CQ882115 (1-2721)			
QY	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
Db	2044	GTGCGCTCCTCCAAGAACGTTCATCAAGGAGTTTCATCGCTTCAAGGTGCGCATGGAGGGC	2103
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40

Alignment Scores:  
Pred. No.: 4.1e-118 Length: 2772



Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservatives: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x CQ882117 (1-2772)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 2095 GTGGCTCCTCCAAGACGTCAACAAGGATTCAATGGCTTCAAGGTGGCATGGAGGGC 2154  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 2155 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC 2214  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 2215 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTGCCTTCGGACATC 2274  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 2275 CTGTCCCCCAGTTCACGTACGGCTCCAAGGTGTACGTGAAGCACCCCGCCGACATCCCC 2334  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
Db 2335 GACTACAAGAAGCTGTCTTCCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTCGAG 2394  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 2395 GACGGCGGCGTGGTGACCGTGACCCAGGACTCCTCCCTGCAGGACGGCTTTCATCTAC 2454  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 2455 AAGGTGAAGTTTCATCGCGTGAACCTTCCCTCCGACGGCCCCCGTAATGCAGAAGAAC 2514  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 2515 ATGGGTGGAGGCTCCACCGAGCGCTGTACCCCGCGACGGCGTGCTGAAGGGCGAG 2574  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 2575 ATCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACCTGGTGGAGTTCAAGTCCATC 2634  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 2635 TACATGGCCCAAGAAGCCCGTCAGCTGCCCGGCTACTACTAGCTGGACTCCAAAGCTGGAC 2694  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 2695 ATCACCTCCCAACGAGGACTACACCATCGTGGAGCAGTACGAGCGCACCGAGGGCCGC 2754  
QY 221 HisHisLeuPheLeu 225  
Db 2755 CACCACCTGTTCTCTG 2769

RESULT 14  
AX463702  
LOCUS AX463702 4692 bp DNA linear PAT 15-JUL-2002  
DEFINITION Sequence 16 from Patent WO0248338.  
ACCESSION AX463702  
VERSION AX463702.1 GI:21886461  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
other sequences; artificial sequences.  
REFERENCE 1  
AUTHORS Lichtenberg-Frat,H.  
TITLE yeast strain for testing the geno- and cytotoxicity of complex environmental contamination  
JOURNAL Patent: WO 0248338-A 16 20-JUN-2002;  
Lichtenberg-Frat, Hella (DE)  
FEATURES Location/Qualifiers

source 1. .4692  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Vektor pDsRed1-N1"

ORIGIN

Alignment Scores:

Pred. No.: 7.96e-118 Length: 4692  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservatives: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x AX463702 (1-4692)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 682 GTGGCTCCTCCAAGAACGTCAACAAGGATTCAATGGCTTCAAGGTGGCATGGAGGGC 741  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 742 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC 801  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 802 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTGCCTTCGGCTGGGACATC 861  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 862 CTGTCCCCCAGTTCACGTACGGCTCCAAGGTGTACGTGAAGCACCCCGCCGACATCCCC 921  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
Db 922 GACTACAAGAAGCTGTCTTCCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTCGAG 981  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 982 GACGGCGGCGTGGTGACCGTGACCCAGGACTCCTCCCTGCAGGACGGCTGCTTCATCTAC 1041  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 1042 AAGGTGAAGTTTCATCGCGTGAACCTTCCCTCCGACGGCCCCCGTAATGCAGAAGAAC 1101  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 1102 ATGGGTGGAGGCTCCACCGAGCGCTGTACCCCGCGACGGCGTGCTGAAGGGCGAG 1161  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 1162 ATCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACCTGGTGGAGTTCAAGTCCATC 1221  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 1222 TACATGGCCCAAGAAGCCCGTGACGTGCCCGGCTACTACTAGCTGGACTCCAAAGCTGGAC 1281  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 1282 ATCACCTCCCAACGAGGACTACACCATCGTGGAGCAGTACGAGCGCACCGAGGGCCGC 1341  
QY 221 HisHisLeuPheLeu 225  
Db 1342 CACCACCTGTTCTCTG 1356

RESULT 15  
AX823860  
LOCUS AX823860 6893 bp DNA linear PAT 11-DEC-2003  
DEFINITION Sequence 8 from Patent WO03070931.  
ACCESSION AX823860  
VERSION AX823860.1 GI:39750176  
KEYWORDS  
SOURCE synthetic construct

ORGANISM	synthetic construct	
REFERENCE	other sequences; artificial sequences.	
1		
AUTHORS	Baum,C., Will,E., Ostertag,W., Klump,H. and Schiedlmeier,B.	
TITLE	Methods for conducting site-specific dna recombination	
JOURNAL	Patent: WO 03070931-A 8 28-AUG-2003;	
	Vision 7 GmbH (DE)	
FEATURES	Location/Qualifiers	
source	1..6893	
	/organism="synthetic construct"	
	/mol_type="unassigned DNA"	
	/db_xref="taxon:32630"	
	/note="Beschreibung der kuenstlichen Sequenz: provirale	
	Plasmid-DNA; retrovirales Cre-Reporterkonstrukt #SFr#"	
	1..435	
misc_feature	/note="Plasmid-Rueckgrat (pUC)"	
misc_feature	436..996	
primer_bind	/note="MPSV 5'-LTR (Deltal-31)"	
	997..1014	
	/note="PBS (primer binding site)"	
5'UTR	1015..1560	
misc_feature	1561..1599	
	/note="#loxP1#-Sequenz"	
misc_feature	1600..2289	
	/note="DsRed1-Gen aus Discosoma sp."	
misc_feature	2290..2343	
	/note="#loxP2#-Sequenz"	
misc_feature	2347..3150	
	/note="eGFP(tag) -Gen aus Aequorea victoria"	
misc_feature	3164..4019	
	/note="#wPRE#"	
3'UTR	4020..4080	
misc_feature	4081..4641	
	/note="3'-LTR"	
misc_feature	4642..6893	
	/note="Plasmid-Rueckgrat (pUC)"	
ORIGIN		
Alignment Scores:		
Pred. No.:	1.29e-117	Length: 6893
Score:	1210.00	Matches: 224
Percent Similarity:	100.00%	Conservative: 1
Best Local Similarity:	99.56%	Mismatches: 0
Query Match:	99.67%	Indels: 0
DB:	6	Gaps: 0
US-10-006-922A-12 (1-225) x AX823860 (1-6893)		
Qy	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
Db	1603	GTGCGCTCCTCCCAAGAACGTCATCAAGGAGTTTCATGCGCTTCAAGGTGCGCATGGAGGGC 1662
Qy	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
Db	1663	ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCTACGAGGGC 1722
Qy	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60
Db	1723	CACAACACCCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTCCCTTCGCCCTGGGACATC 1782
Qy	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80
Db	1783	CTGTCCCCCAGTTCCAGTAGCGGTCCAAGGTGTACGTGAAGCACCCCGCGACATCCCC 1842
Qy	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100
Db	1843	GACTACAAGAAGCTGTCTTCCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTCGAG 1902
Qy	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
Db	1903	GACGGCGGGCGTGGTGACCGTGACCCAGGACTCTCCCTGCAGGACGGGTGCTTCATCTAC 1962
Qy	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140

Db	1963	AAGGTGAAGTTTCATCGGCGTGAACCTTCCCTCCGACGGCCCCGTAATGCAGAAGACC 2022
Qy	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
Db	2023	ATGGGCTGGAGGGCCTCCACCGAGCGCCTGTACCCCGCGACGGCGTGCTGAAGGGCGAG 2082
Qy	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
Db	2083	ATCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACCTGGTGGAGTTCAAGTCCATC 2142
Qy	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
Db	2143	TACATGGCCAAGAAGCCCCGTGCAGCTGCCCGGTACTACTACTGGACTCCAAGCTGGAC 2202
Qy	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
Db	2203	ATCACCTCCACACAACGAGGACTACACCATCGTGGAGCAGTACGAGCGCACCGAGGGCCGC 2262
Qy	221	HisHisLeuPheLeu 225
Db	2263	CACCACCTGTTTCCTG 2277

Search completed: July 1, 2005, 19:55:32  
Job time : 4404 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 15:39:59 ; Search time 40 Seconds  
(without alignments)  
541.219 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVIKEFMRFKVRMEG.....EDYTIVEQYERTEGRHHLFL 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	211.5	17.4	238	1 JQ1514	green-fluorescent
2	93.5	7.7	26926	1 I38344	titin, cardiac mus
3	92	7.6	1873	2 T30944	surface protein pr
4	91	7.5	458	1 A42386	hsp 90-binding pro
5	89	7.3	1433	1 A36734	bacillopeptidase F
6	87.5	7.2	1116	2 B70476	hypothetical prote
7	87	7.2	568	2 T06489	probable peptidylp
8	87	7.2	1484	2 C97196	probable membrane
9	85.5	7.0	456	2 G69397	signal-transducing
10	85	7.0	271	2 F69442	hypothetical prote
11	84.5	7.0	679	2 A40351	adhesion-type prot
12	84.5	7.0	725	1 IJMSNG	neural cell adhesi
13	84	6.9	1616	2 T17884	S-layer protein -
14	83.5	6.9	268	2 E90276	conserved hypothet
15	83.5	6.9	340	2 E69544	hypothetical prote
16	83.5	6.9	374	2 T06245	gibberellin 3 beta
17	83.5	6.9	374	2 T06244	gibberellin 3 beta
18	83	6.8	373	2 T50605	hypothetical prote
19	83	6.8	15281	2 S41309	cyclosporin synthe
20	82.5	6.8	559	1 S55383	peptidylprolyl iso
21	82.5	6.8	680	2 S17982	Kallmann syndrome
22	82.5	6.8	931	2 T32919	hypothetical prote
23	82	6.8	341	2 B53125	restriction enzyme
24	81.5	6.7	292	2 C69106	glucose-1-phosphat
25	81.5	6.7	296	2 I37989	La 4.1 protein - h
26	81.5	6.7	551	1 S72485	peptidylprolyl iso
27	81.5	6.7	629	2 C64180	hypothetical prote
28	81.5	6.7	6805	2 S20901	titin - rabbit (fr
29	81	6.7	346	2 S77025	nitrilase (EC 3.5.

30 81 6.7 587 1 E69171 phosphoesterase-re  
31 80.5 6.6 404 1 S03849 ribonucleoprotein  
32 80.5 6.6 415 1 JC1494 ribonucleoprotein  
33 80.5 6.6 862 2 F75116 hypothetical prote  
34 80.5 6.6 1003 1 AJCHPR phosphoribosylamin  
35 80.5 6.6 1115 1 IJMSNL neural cell adhesi  
36 80 6.6 621 2 A95250 choline-binding pr  
37 80 6.6 690 2 F98114 choline-binding pr  
38 79.5 6.5 459 2 A46372 immunophilin FKBP5  
39 79.5 6.5 475 2 T44566 conserved hypothet  
40 79.5 6.5 487 2 T45982 hypothetical prote  
41 79.5 6.5 648 2 B84139 ABC transporter (p  
42 79.5 6.5 853 1 IJBONC neural cell adhesi  
43 79 6.5 458 2 JN0873 immunophilin p59 -  
44 79 6.5 703 2 A64351 hypothetical prote  
45 79 6.5 1214 2 T47438 disease resistance

ALIGNMENTS

RESULT 1

JQ1514  
green-fluorescent protein [validated] - hydromedusa (Aequorea victoria)  
C;Species: Aequorea victoria  
C;Date: 03-Dec-1999 #sequence revision 03-Dec-1999 #text change 09-Jul-2004  
C;Accession: JS0692; JQ1514; FQ0335; S48693; S51330; S51331  
R;Prasher, D.C.; Eckenrode, V.K.; Ward, W.W.; Prendergast, F.G.; Cormier, M.J.  
Gene 111, 229-233, 1992  
A;Title: Primary structure of the Aequorea victoria green-fluorescent protein.  
A;Reference number: JQ1514; MUID:92175527; PMID:1347277  
A;Accession: JS0692  
A;Molecule type: DNA  
A;Residues: 1-107,'S',109-238 <PRA1>  
A;Cross-references: UNIPROT:P42212; UNIPROT:Q17106; UNIPROT:Q17105; GB:M62654; NID:g1556  
A;Accession: JQ1514  
A;Molecule type: mRNA  
A;Residues: 1-99,'F',101-140,'L',142-218,'V',220-238 <PRA2>  
A;Cross-references: GB:M62653; NID:g155660; PIDN:AAA27721.1; PID:g155661  
A;Accession: PQ0335  
A;Molecule type: protein  
A;Residues: 46-64;74-122;132-151;154-183;185-200 <PRA3>  
R;Inouye, S.; Tsuji, F.I.  
FEBS Lett. 351, 211-214, 1994  
A;Title: Evidence for redox forms of the Aequorea green fluorescent protein.  
A;Reference number: S48693; MUID:94364470; PMID:8082767  
A;Accession: S48693  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-24,'Q',26-156,'P',158-171,'K',173-238 <INO>  
A;Cross-references: GB:L29345; NID:g606383; PIDN:AAA58246.1; PID:g606384  
R;Watkins, J.N.; Campbell, A.K.  
Submitted to the EMBL Data Library, January 1995  
A;Reference number: S51330  
A;Accession: S51330  
A;Molecule type: mRNA  
A;Residues: 1-13,'V',15-24,'Q',26-44,'N',46-153,'G',155-156,'P',158-171,'K',173-227,'R',  
A;Cross-references: EMBL:X83959; NID:g634008; PIDN:CAA58789.1; PID:g634009  
A;Experimental source: clone gfp1  
A;Accession: S51331  
A;Molecule type: mRNA  
A;Residues: 1-24,'Q',26-29,'R',31-83,'L',85-153,'G',155-156,'P',158-171,'K',173-208,'Q',  
A;Cross-references: EMBL:X83960; NID:g634010; PIDN:CAA58790.1; PID:g634011  
A;Experimental source: clone gfp2  
R;Yang, F.; Moss, L.G.; Phillips Jr., G.N.  
submitted to the Brookhaven Protein Data Bank, August 1996  
A;Reference number: A65692; PDB:1GFL  
A;Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 'A',2-79,'R',81-9  
A;Note: engineered sequence based on JQ1514, cloned and expressed in Escherichia coli  
R;Yang, F.; Moss, L.G.; Phillips Jr., G.N.  
Nat. Biotechnol. 14, 1246-1251, 1996  
A;Title: The molecular structure of green fluorescent protein.  
A;Reference number: A58953; MUID:98294543; PMID:9631087







A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serod  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpsira, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A;Reference number: A69580; MUID:98044033; PMID:9384377  
A;Accession: B69596  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-1433 <KUN>  
A;Cross-references: GB:Z99111; GB:Z99112; GB:AL009126; NID:g2633902; PIDN:CAB13404.1; PI  
A;Experimental source: strain 168  
C;Genetics:  
A;Gene: bpr; bpf  
A;Map position: 135 (degrees)  
C;Superfamily: bacillopeptidase F; subtilisin homology  
C;Keywords: extracellular protein; hydrolase; serine proteinase  
F;1-30/Domain: signal sequence #status predicted <SIG>  
F;31-194/Domain: propeptide #status predicted <PRO>  
F;195-1433/Product: bacillopeptidase F #status experimental <MAT>  
F;218-466/Domain: subtilisin homology <SBT>  
F;227,274,452/Active site: Asp, His, Ser #status predicted

Query Match 7.3%; Score 89; DB 1; Length 1433;  
Best Local Similarity 24.0%; Pred. No. 22;  
Matches 56; Conservative 26; Mismatches 71; Indels 80; Gaps 14;

QY 17 RMEGTVNGHEFEIEGEGRPYEGHNTVKLVKVTGGPLPFA----- 57  
Db 521 KAEGQVS-----VEGDDQEPVPVYQHEKV-TEAYEGGSLPTLTAEDNVSVTSVKLSYKLD 574  
QY 58 ---WDILSPQFOYGSKVYVKHPADIPDYK--KLSFPPEGFKWERVMNFEDEG-----VVT 106  
Db 575 QGEWTEITAKRISGDHLKGTQAEIPDIKGTKLS---YKW---MIHDRGHWVSSDVYD 627  
QY 107 VTQDSSLQDGCFCFIYVKVF-----IGVN-----FPSDGPVMQKKTMGWEASTERLY 151  
Db 628 VTVKPSITAG---YKQDFETAPGGWVASGTNNWNGWVPSTGP-----NTAASGEKVY 677  
QY 152 PRDGVLKGE-----IHKALKLKDGGHYLVEFKSIYMAKKPVQLPGYVV 195  
Db 678 GTN--LTGNYANSANNLVMPPIKAPDSGSLFLQPKSWHNLEDDFDY-GYFV 727

RESULT 6  
B70476  
hypothetical protein aq\_2054 - Aquifex aeolicus  
C;Species: Aquifex aeolicus  
C;Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004  
C;Accession: B70476  
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O  
V.  
Nature 392, 353-358, 1998  
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A;Reference number: A70300; MUID:98196666; PMID:9537320  
A;Accession: B70476  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-1116 <AQF>  
A;Cross-references: UNIPROT:O67838; GB:AE000770; NID:g2984274; PIDN:AAC07805.1; PID:g298  
A;Experimental source: strain VF5  
C;Genetics:  
A;Gene: aq\_2054

Query Match 7.2%; Score 87.5; DB 2; Length 1116;  
Best Local Similarity 22.0%; Pred. No. 21;  
Matches 56; Conservative 35; Mismatches 103; Indels 61; Gaps 9;

QY 5 KNVIKEFMRFKVRMEGTVNGHEFEIEGEGE-----GRPYEGHNTVKLVKVTGG 52  
Db 164 KOLLDDSEYFSAKTKGEIKRNTGEILAEVEIKEIRKENFTLSGTKINAKGTINLPVLDIN 223  
QY 53 PLPFAWDILSPQFQYGS-----KVYVKHPADIPDYKKLSF-----PEG---F 91

Db 224 AKAFVKDLIVRNKNYGSIEGIVKGNVELFDKFLKGEAVNPEGTIKIKFTYDVIPEGLLTF 283  
QY 92 KWER-VNMFEDGGVTVVTQDSSLQDGCFCFIYKVKFIGVNFPSDGPVMQKKTMGWEASTERL 150  
Db 284 SFENLVVDKNTLGINREIRGEFHGNGKVDFKMFVKVN-----AFTENL 327  
QY 151 YPRDGVLKGEIHKALKLKDGGHYLVEFKSIYMAKKPVQLPGYVVYVD-SKLDITSHNEDYT 209  
Db 328 EVIDKKFKGDVLFSYNFSNGSLNFEFKNSGYAK-----GNLIINKNKLEGEFSFNDFP 381  
QY 210 IVEQ-----YERTEGR 220  
Db 382 VVFQDFNAYLSGEGK 396

RESULT 7  
T06489  
probable peptidylprolyl isomerase (EC 5.2.1.8) FKBP77 - wheat  
C;Species: Triticum aestivum (common wheat)  
C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C;Accession: T06489  
R;Brieman, A.  
submitted to the EMBL Data Library, August 1996  
A;Reference number: Z15713  
A;Accession: T06489  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-568 <BRI>  
A;Cross-references: UNIPROT:O04843; EMBL:Y07636; PIDN:CAA68913.1  
A;Experimental source: cv. ATIR, 2 day old plants, root tips  
C;Genetics:  
A;Gene: FKBP77  
C;Superfamily: peptidylprolyl isomerase ROF1; BKBP-type peptidylprolyl isomerase homology  
C;Keywords: calmodulin binding; cis-trans-isomerase; cyclosporin A binding  
F;54-101/Domain: BKBP-type peptidylprolyl isomerase homology <PPI>  
F;169-212/Domain: BKBP-type peptidylprolyl isomerase homology <PPI2>  
F;286-334/Domain: BKBP-type peptidylprolyl isomerase homology <PPI1>  
F;482-515/Domain: tetratricopeptide repeat homology <TTR>

Query Match 7.2%; Score 87; DB 2; Length 568;  
Best Local Similarity 21.6%; Pred. No. 9.6;  
Matches 61; Conservative 32; Mismatches 81; Indels 108; Gaps 15;

QY 5 KNVIKE-----FMRFKVRMEGT-VNGHEFEIEGEGRPYEGHNTVKLVKVTGGPL 54  
Db 39 KKLVKEGEGWDTAETALKVEVHYTGTLDDGTKFD-SSRDRGTPTF-----KFKLEQGQVI 91  
QY 55 PFAWD-----ILSPQFQYGSKVYVKHPADIPDYKKLSFP----- 88  
Db 92 K-GWDQGIKTMKKGNASLTIPPDLAYGERA---PRTIPPNTALRFDVELLSWASVKDI 146  
QY 89 -----EGFKWERVMNFEDEGGVTVVTQDSSLQDGCFCFIYKVKFIGVNFPSDGPVM 136  
Db 147 CKDGGIFKKVLVEGQKWE---NPXDLDEVTVKYEARLEDGGSVVSKSESI----- 192  
QY 137 QKKTMGWEASTERLYPRDGVLKGEIHKALK-LKDGGHYLVEFKSIY---MAKKPVQLPG 191  
Db 193 -----EFSV-----KQGYFCPALSKAVKTMKGEKVLTLTVKBPQYGFGEQGRAATEVEG 240  
QY 192 YYYVDSKLDI-----TSHNEDYTIV-----EQYER 216  
Db 241 AVPPNSTLHIDLQLVSWKTLTLIGDDKRILKVKLKEGEGYER 282

RESULT 8  
C97196  
probable membrane protein [imported] - Clostridium acetobutylicum  
C;Species: Clostridium acetobutylicum  
C;Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 09-Jul-2004  
C;Accession: C97196  
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.





Qy	91	FKWERMNFEDGGVVTVTQDSSLQDGCFIYKVKFIGVNFPSDGPVMQKKTMGWEASTERL	150
Db	466	CAHNRTTGSEASSGWMHENYIILQDLSFSCKYK-----TVQPIRPKS-----HSKAEAV	515
Qy	151	Y---PRDGLVKGEIHKALK-LKDGGHYLVEFKSIYMAKKPVQLPGYVVYVDSKLDITSH	204
Db	516	FFTTPPCSALKGSHKPIGCLGEAGHVLSK-----VLAKPENLSASFIVQD-VNITGH	567
RESULT 12			
IJMSNG			
neural cell adhesion molecule 1 precursor, GPI-anchored splice form - mouse			
N;Alternate names: NCAM-120			
C;Species: Mus musculus (house mouse)			
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 09-Jul-2004			
C;Accession: A29673; S00382; A44290			
R;Barthels, D.; Santoni, M.J.; Wille, W.; Ruppert, C.; Chaix, J.C.; Hirsch, M.R.; Fonted			
EMBO J. 6, 907-914, 1987			
A;Title: Isolation and nucleotide sequence of mouse NCAM cDNA that codes for a Mr 79,000			
A;Reference number: A29673; MUID:87246524; PMID:3595563			
A;Accession: A29673			
A;Molecule type: mRNA			
A;Residues: 1-725 <BAR>			
A;Cross-references: UNIPROT:P13594; EMBL:Y00051; NID:G53342; PIDN:CAA68263.1; PID:G53343			
R;Barbas, J.A.; Chaix, J.C.; Steinmetz, M.; Goridis, C.			
EMBO J. 7, 625-632, 1988			
A;Title: Differential splicing and alternative polyadenylation generates distinct NCAM b			
A;Reference number: S00382; MUID:88283628; PMID:3396534			
A;Accession: S00382			
A;Molecule type: DNA			
A;Residues: 642-656, 'D', 658-725 <BA2>			
A;Cross-references: EMBL:X07195			
R;Rougon, G.; Marshak, D.R.			
J. Biol. Chem. 261, 3396-3401, 1986			
A;Title: Structural and immunological characterization of the amino-terminal domain of m			
A;Reference number: A44290; MUID:86140120; PMID:3512556			
A;Accession: A44290			
A;Molecule type: protein			
A;Residues: 20-36 <ROU>			
C;Comment: NCAM mediates cell-cell adhesion via homophilic binding with another NCAM mol			
C;Comment: Several forms of NCAM are produced by alternative splicing. See also PIR:IJMS			
C;Genetics:			
A;Gene: NCAM			
A;Map position: 9			
A;Introns: 701/1			
C;Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; immu			
C;Keywords: alternative splicing; cell adhesion; duplication; heparin binding; membrane			
F;1-19/Domain: signal sequence #status predicted <SIG>			
F;34-98/Domain: immunoglobulin homology <IMM1>			
F;132-191/Domain: immunoglobulin homology <IMM2>			
F;152-156/Region: heparin binding #status predicted			
F;161-165/Region: heparin binding #status predicted			
F;228-290/Domain: immunoglobulin homology <IMM3>			
F;263-272/Region: NCAM binding #status predicted			
F;323-388/Domain: immunoglobulin homology <IMM4>			
F;420-482/Domain: immunoglobulin homology <IMM5>			
F;519-596/Domain: fibronectin type III repeat homology <FN3A>			
F;625-685/Domain: fibronectin type III repeat homology <FN3B>			
F;41-96,139-189,235-288,330-386,427-480/Disulfide bonds: #status predicted			
F;222,316,348,424,450,479/Binding site: carbohydrate (Asn) (covalent) #status predicted			
Query Match			
Best Local Similarity 7.0%; Score 84.5; DB 1; Length 725;			
Matches 37; Conservative 30; Mismatches 57; Indels 29; Gaps 9;			
Qy	91	FKWERMNFEDGGVVTV---TQDSSLQDGCFIYKVKFIGVNFPSDGPVMQKKTMGWEAST	147
Db	548	FTWYDAKEANMEGIVTIMGLKPETTSYD-----RLAALNGKGLGEIMQPSKTPQVP	600
Qy	148	ERLYPRDGLVKGEIHK-----ALKLKDGG---HYLVEFKSIYMAKKP-VQLP-GYY	193
Db	601	ELSAPK---LEGQMGEDGNSIKWNLIKQDDGGSPRHYLVKYRALASEWKPFIRLPSGSH	657

Qy	194	YVDSK-LDITSHNEDYTIVE-QYERTEGRHHLF	224
Db	658	HVMLKSLDWNAEYEVYVVAENQOGKSKAAHFVF	690
RESULT 13			
T17884			
S-layer protein - Bacillus circulans			
C;Species: Bacillus circulans			
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004			
C;Accession: T17884			
R;Aubert-Pivert, E.; Davies, J.			
Gene 147, 1-11, 1994			
A;Title: Biosynthesis of butirosin in Bacillus circulans NRRL B3312: identification by se			
A;Reference number: Z18808; MUID:94374689; PMID:7522196			
A;Accession: T17884			
A;Status: preliminary; translated from GB/EMBL/DBDJ			
A;Molecule type: DNA			
A;Residues: 1-1616 <AUB>			
A;Cross-references: UNIPROT:P35824; EMBL:L20421; NID:G304142; PID:G304143; PIDN:AAA62588			
C;Genetics:			
A;Gene: butB			
C;Function:			
A;Pathway: butirosin biosynthesis			
Query Match			
Best Local Similarity 6.9%; Score 84; DB 2; Length 1616;			
Matches 59; Conservative 48; Mismatches 85; Indels 94; Gaps 17;			
Qy	6	NVIKEFMR-----FKVRMEGTVNGHEFEI---EGEGGRPYEGHNTVKLKVTK	50
Db	590	DLLQEFIRYSRELGLDIHVSFNIFAEGSIASNEFALLDSDLWBEERVYNAADNGQIK---	646
Qy	51	GGPLPFAWDILSPQFYQYGSKVYVKHPAD-IPDYKKLSFPEGFKWERVMNFEDGGVV--TV	107
Db	647	-----RLRESAKQGAFAFVNPSNDEVDRDFQLKTIE-----EVLQNYDVDGVLDRA	692
Qy	108	TQDSSLQDGCFIYKVK---FIGV-----NFP-----SDGPVMQ-----KKT	140
Db	693	RYDNESADFSDLTKAKFESFLGARGKQLQNWPDVFTYAGNVRKDGPLIRDWWEFRSKTI	752
Qy	141	MGWEASTERLYPRDGLVKGEIHKALKLK-----DGGHY-LVEFK---SIYMAK	184
Db	753	KSFTSEVRQLTDR---VKAEEKGKKIEVSAYVGSWFESYYLNGVHWGSTEFTRYDERLRMKD	809
Qy	185	KPVQLPGYY---YVDSKLDITSHNEDYTIVEQYERT--EGRHHLFL	225
Db	810	KSVYTPGYYESGYV-----KNLDPFIMIGAYQTTAPEIEHYITL	847
RESULT 14			
E90276			
conserved hypothetical protein [imported] - Sulfolobus solfataricus			
C;Species: Sulfolobus solfataricus			
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004			
C;Accession: E90276			
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-v			
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.			
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.			
submitted to GenBank, April 2001			
A;Description: Sulfolobus solfataricus complete genome.			
A;Reference number: A99139			
A;Accession: E90276			
A;Status: preliminary			
A;Molecule type: DNA			
A;Residues: 1-268 <KUR>			
A;Cross-references: UNIPROT:Q97YT6; GB:AE006641; NID:G13814420; PIDN:AAK41468.1; GSPDB:GN			
C;Genetics:			
A;Gene: SSO1221			
Query Match			
Best Local Similarity 6.9%; Score 83.5; DB 2; Length 268;			
Matches 49; Conservative 21; Mismatches 59; Indels 61; Gaps 13;			

Qy 34 EGRPYEGHNTVKLVTKGG-----PLPFAWDILSPQFYGSKVY-----VKHPAD 78  
| | | | : | | | : | | | : | | | : | | | :  
Db 3 EIRPY--YNTFSLPITKSGKSQIVPPPPWIYAI-----EMIGVKAFFDPVKVLDLVPPPLE 56  
| | | | : | | | : | | | : | | | : | | | :  
Qy 79 IPD-----YKLSFPEGFKWERVMNFEDGGVVVTQDSSLQDGCIFYKVKF---IGVNF 129  
| | | | : | | | : | | | : | | | : | | | :  
Db 57 IVDGEGFVYIAKIFTVSGNRWE--MLYED-----PEETKYMEAAIALKVKNYDNIFTYF 108  
| | | | : | | | : | | | : | | | : | | | :  
Qy 130 P-----SDGPVMQKKTMGWEASTERLYPRD--GVLKGEIHKALK-----LKDGGHYL 174  
| | | | : | | | : | | | : | | | : | | | :  
Db 109 PFMWVDKOLPLLR-----GWLLG----YPKLAYISISEFHKLLDGYSGPSSGVRMGGYAL 160  
| | | | : | | | : | | | : | | | : | | | :  
Qy 175 VEFKSIYMAK 184  
| | | | : | | | : | | | : | | | : | | | :  
Db 161 RNGKEIIRVK 170  
| | | | : | | | : | | | : | | | : | | | :  
  
RESULT 15  
E69544  
hypothetical protein AF2357 - Archaeoglobus fulgidus  
C;Species: Archaeoglobus fulgidus  
C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C;Accession: E69544  
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.  
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
Nature 390, 364-370, 1997  
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.  
Smith, H.O.; Woese, C.R.; Venter, J.C.  
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo  
A;Reference number: A69250; MUID:98049343; PMID:9389475  
A;Accession: E69544  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-340 <KLE>  
A;Cross-references: UNIPROT:O30313; GB:AE001113; GB:AE000782; NID:g2689436; PIDN:AAB9130  
  
Query Match 6.9%; Score 83.5; DB 2; Length 340;  
Best local Similarity 23.9%; Pred. No. 9.9;  
Matches 38; Conservative 29; Mismatches 63; Indels 29; Gaps 9;  
  
Qy 41 HNTVKLVTKGGPLPFAWDILSPQFYGSKVYVVKHPADIPDYKKLSFPEGFKWERVMNFE 100  
| | | | : | | | : | | | : | | | : | | | :  
Db 144 HAWVEVKINN-----SWVADPTVYW---LYVNYPEKYPNWNKLPF--NNESSWANLIDF- 192  
| | | | : | | | : | | | : | | | : | | | :  
Qy 101 DGGVVTVTQDSSLQDGCIFYKVKFIGVNFPSDGPVMQ--KKTMGWEASTERLYPRDGVLK 158  
| | | | : | | | : | | | : | | | : | | | :  
Db 193 -SRVTVLPNGSVLDLTSNY-TKTYNVTTITIDQNVKGIKVTTWKGSVERTVYSKAVNK 250  
| | | | : | | | : | | | : | | | : | | | :  
Qy 159 GE-IHKALKLDGGHYLVEFKSIYMAKKPVQLPGYYVVD 196  
| | | | : | | | : | | | : | | | : | | | :  
Db 251 SDTVNLALATR-----IYKFELI-----VPTWYFLE 276  
| | | | : | | | : | | | : | | | : | | | :

Search completed: June 30, 2005, 15:50:22  
Job time : 42 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 15:31:48 ; Search time 175 Seconds  
(without alignments)  
658.388 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVIKEFMRFKVRMEG.....EDYTIVEQYERTEGRHHLPL 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1214	100.0	225	2	Q9U6Y8	Q9u6y8 discosoma s
2	1196	98.5	236	2	Q66ND7	Q66nd7 discosoma s
3	1193	98.3	225	2	Q6KF85	Q6kff85 discosoma s
4	1191	98.1	236	2	Q66ND8	Q66nd8 discosoma s
5	1085.5	89.4	230	2	Q9GTJ7	Q9gtj7 discosoma s
6	804	66.2	221	2	Q66PV1	Q66pv1 acropora te
7	798	65.7	221	2	Q66PV0	Q66pv0 acropora mi
8	797	65.7	221	2	Q66PV0	Q66pv0 acropora mi
9	793	65.3	221	2	Q95P04	Q95p04 gonopora t
10	793	65.3	221	2	Q66PV9	Q66pv9 acropora hy
11	751	61.9	219	2	P83690	P83690 montipora e
12	745	61.4	227	2	Q66ND6	Q66nd6 discosoma s
13	729.5	60.1	232	2	Q9U6Y7	Q9u6y7 discosoma s
14	690	56.8	225	2	Q963F5	Q963f5 montastraea
15	684	56.3	225	2	Q66ND3	Q66nd3 montastraea
16	681	56.1	225	2	Q95UA7	Q95ua7 montastraea
17	681	56.1	225	2	Q7Z0W4	Q7z0w4 montastraea
18	679.5	56.0	227	2	Q7Z0W6	Q7z0w6 montastraea
19	678.5	55.9	227	2	Q962P9	Q962p9 montastraea
20	678.5	55.9	227	2	Q7Z0W8	Q7z0w8 montastraea
21	678	55.8	225	2	Q7Z0W5	Q7z0w5 montastraea
22	672.5	55.4	227	2	Q66ND2	Q66nd2 montastraea
23	670.5	55.2	227	2	Q66ND5	Q66nd5 montastraea
24	669.5	55.1	234	2	Q7Z0W7	Q7z0w7 montastraea
25	668.5	55.1	234	2	Q8T5F2	Q8t5f2 montastraea
26	667	54.9	266	2	Q9U6Y3	Q9u6y3 clavularia
27	663	54.6	226	2	Q8T6U0	Q8t6u0 dendroneph
28	660.5	54.4	227	2	Q66ND4	Q66nd4 montastraea
29	655.5	54.0	226	2	Q66PV5	Q66pv5 agaricia fr
30	655.5	54.0	227	2	Q95VT0	Q95vt0 montastraea
31	654.5	53.9	234	2	Q8MU47	Q8mu47 montastraea

32	652.5	53.7	225	2	Q7Z0W9	Q7z0w9 montastraea
33	642.5	52.9	225	2	Q6USK3	Q6usk3 montastraea
34	639	52.6	224	2	Q8MU48	Q8mu48 montastraea
35	631.5	52.0	225	2	Q8T5F1	Q8t5f1 montastraea
36	620.5	51.1	230	2	Q66PW1	Q66pw1 scolymia cu
37	610	50.2	223	2	Q6R8F5	Q6r8f5 astrangia l
38	608.5	50.1	231	2	Q8ISF8	Q8isf8 parasicyoni
39	599.5	49.4	225	2	Q8I6J8	Q8i6j8 trachyphyll
40	580.5	47.8	219	2	Q6R8F4	Q6r8f4 astrangia l
41	579.5	47.7	259	2	Q8MMA2	Q8mma2 agaricia fr
42	574	47.3	228	2	Q9GPI6	Q9gpi6 anemonia su
43	573.5	47.2	235	2	Q8T5F0	Q8t5f0 scolymia cu
44	570.5	47.0	231	2	Q8T6T8	Q8t6t8 discosoma s
45	566.5	46.7	239	2	Q8MMA1	Q8mma1 agaricia ag

ALIGNMENTS

RESULT 1

Q9U6Y8 PRELIMINARY; PRT; 225 AA.  
AC Q9U6Y8;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Fluorescent protein FP583.  
OS Discosoma sp.  
OC Eukaryota; Metazoa; Chnidaria; Anthozoa; Zoantharia; Corallimorpharia;  
OC Discosomatidae; Discosoma.  
OX NCBI\_TaxID=86600;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99436614; PubMed=10504696;  
RA Matz M.V., Fradkov A.F., Labas Y.A., Savitsky A.P., Zairaisky A.G.,  
RA Markelov M.L., Lukyanov S.A.;  
RT "Fluorescent proteins from nonbioluminescent Anthozoa species.";  
RL Nat. Biotechnol. 17:969-973(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Matz M.V., Fradkov A.F., Labas Y.A., Savitsky A.P., Zairaisky A.G.,  
RA Markelov M.L., Lukyanov S.A.;  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF168419; AAF03369.1; -  
DR PDB; 1G7K; X-ray; A/B/C/D=1-225.  
DR PDB; 1GGX; X-ray; A/B/C/D=1-225.  
DR GO; GO:0006091; P:energy pathways; IEA.  
DR InterPro; IPR009017; GFP\_like.  
DR InterPro; IPR011584; GFP\_related.  
DR InterPro; IPR000786; Green\_fl\_protein.  
DR Pfam; PF01353; GFP; 1.  
DR PRINTS; PR01229; GFP-related.  
DR ProDom; PD013756; Green\_fl\_protein; 1.  
SQ SEQUENCE 225 AA; 25931 MW; FBF9AS369778F689 CRC64;

Query Match 100.0%; Score 1214; DB 2; Length 225;  
Best Local Similarity 100.0%; Pred. No. 2.9e-98;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVTKGGPLPFAWDI	60
Db	1	MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVTKGGPLPFAWDI	60
QY	61	LSPQFQYGSKVYVVKHPADIPDYKKLSFPFGFKWERMVNFEDGGVVTVTQDSSLDGCFIY	120
Db	61	LSPQFQYGSKVYVVKHPADIPDYKKLSFPFGFKWERMVNFEDGGVVTVTQDSSLDGCFIY	120
QY	121	KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKDGGHYLVEFKSI	180
Db	121	KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKDGGHYLVEFKSI	180
QY	181	YMAKPVQLPGYVYVDSKLDITSHNEDYIVEQYERTEGRHHLFL	225

Db 181 YMAKKPVQLPGYVVVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225

RESULT 2

Q66ND7 PRELIMINARY; PRT; 236 AA.

ID Q66ND7; Score 1196; DB 2; Length 236; Best Local Similarity 98.2%; Pred. No. 1.2e-96; Matches 221; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

DT 25-OCT-2004 (TrEMBLrel. 28, Created)

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Enhanced red fluorescent protein R+.

OS Discosoma sp. RC-2004.

OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;

OC Discosomatidae; Discosoma.

OX NCBI\_TaxID=289055;

RN [1]

RP SEQUENCE FROM N.A.

RA Carter R.W., Gibbs P.D.L., Schmale M.C.;

RT "Cloning of Anthozoan Fluorescent Protein Genes.";

RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY679107; AAU0444.1; -.

DR InterPro; IPR009017; GFP\_like.

DR InterPro; IPR011584; GFP\_related.

DR InterPro; IPR000786; Green\_fl\_protein.

DR Pfam; PF01353; GFP; 1.

DR PRINTS; PR01229; GFLUORESCENT.

DR ProDom; PD013756; Green\_fl\_protein; 1.

SQ SEQUENCE 236 AA; 27032 MW; BB3844BE06829EF0 CRC64;

Query Match 98.5%; Score 1196; DB 2; Length 236; Best Local Similarity 98.2%; Pred. No. 1.2e-96; Matches 221; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPEYEGHNTVKLVTKGGPLPFAWDI 60

Db 1 MSCSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPEYEGHNTVKLVTKGGPLPFAWDI 60

QY 61 LSPQFQYGSKVYVVKHPADIPDYKKLSFPEGFKWERMNFEDGGVVTVTQDSSLQDGCIFY 120

Db 61 LSPQFQYGSKVYVVKHPADIPDYKKLSFPEGFKWERMNFEDGGVVTVTQDSSLQDGCIFY 120

QY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLVKG EIHKALKLDGGHYLVEFKSI 180

Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLVKG EIHKALKLDGGHYLVEFKTI 180

QY 181 YMAKKPVQLPGYVVVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225

Db 181 YMAKKPVQLPGYVVVDSKLDITSHNKDYTYVEQYERTEGRHHLFL 225

RESULT 3

Q6KF85 PRELIMINARY; PRT; 225 AA.

ID Q6KF85; Score 1191; DB 2; Length 236; Best Local Similarity 97.8%; Pred. No. 3.2e-96; Matches 220; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Orange fluorescent protein FP586.

OS Discosoma sp. JW-2002.

OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;

OC Discosomatidae; Discosoma.

OX NCBI\_TaxID=208461;

RN [1]

RP SEQUENCE FROM N.A.

RA Wiedenmann J., Girod A.;

RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF545828; AAQ11987.1; -.

DR HSSP; P42212; 1B9C.

DR GO; GO:0006091; P:energy pathways; IEA.

DR InterPro; IPR009017; GFP\_like.

DR InterPro; IPR011584; GFP\_related.

DR InterPro; IPR000786; Green\_fl\_protein.

DR Pfam; PF01353; GFP; 1.

DR PRINTS; PR01229; GFLUORESCENT.

DR ProDom; PD013756; Green\_fl\_protein; 1.

SQ SEQUENCE 225 AA; 25791 MW; E151D0E497AA23FA CRC64;

Query Match 98.3%; Score 1193; DB 2; Length 225; Best Local Similarity 98.2%; Pred. No. 2e-96; Matches 221; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPEYEGHNTVKLVTKGGPLPFAWDI 60

Db 1 MSCSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPEYEGHNTVKLVTKGGPLPFAWDI 60

QY 61 LSPQFQYGSKVYVVKHPADIPDYKKLSFPEGFKWERMNFEDGGVVTVTQDSSLQDGCIFY 120

Db 61 LSPQFQYGSKVYVVKHPADIPDYKKLSFPEGFKWERMNFEDGGVVTVTQDSSLQDGCIFY 120

QY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLVKG EIHKALKLDGGHYLVEFKSI 180

Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLVKG EIHKALKLDGGHYLVEFKSI 180

QY 181 YMAKKPVQLPGYVVVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225

Db 181 YMAKKPVQLPGYVVVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225

RESULT 4

Q66ND8 PRELIMINARY; PRT; 236 AA.

ID Q66ND8; Score 1191; DB 2; Length 236; Best Local Similarity 97.8%; Pred. No. 3.2e-96; Matches 220; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

DT 25-OCT-2004 (TrEMBLrel. 28, Created)

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Red fluorescent protein R1.

OS Discosoma sp. RC-2004.

OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;

OC Discosomatidae; Discosoma.

OX NCBI\_TaxID=289055;

RN [1]

RP SEQUENCE FROM N.A.

RA Carter R.W., Gibbs P.D.L., Schmale M.C.;

RT "Cloning of Anthozoan Fluorescent Protein Genes.";

RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY679106; AAU0444.1; -.

DR InterPro; IPR009017; GFP\_like.

DR InterPro; IPR011584; GFP\_related.

DR InterPro; IPR000786; Green\_fl\_protein.

DR Pfam; PF01353; GFP; 1.

DR PRINTS; PR01229; GFLUORESCENT.

DR ProDom; PD013756; Green\_fl\_protein; 1.

SQ SEQUENCE 236 AA; 27042 MW; BB3844BCD6834EF3 CRC64;

Query Match 98.1%; Score 1191; DB 2; Length 236; Best Local Similarity 97.8%; Pred. No. 3.2e-96; Matches 220; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPEYEGHNTVKLVTKGGPLPFAWDI 60

Db 1 MSCSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPEYEGHNTVKLVTKGGPLPFAWDI 60

QY 61 LSPQFQYGSKVYVVKHPADIPDYKKLSFPEGFKWERMNFEDGGVVTVTQDSSLQDGCIFY 120

Db 61 LSPQFQYGSKVYVVKHPADIPDYKKLSFPEGFKWERMNFEDGGVVTVTQDSSLQDGCIFY 120

QY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLVKG EIHKALKLDGGHYLVEFKSI 180

Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLVKG EIHKALKLDGGHYLVEFKTI 180

QY 181 YMAKKPVQLPGYVVVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225

Db 181 YMAKKPVQLPGYVVVDSKLDITSHNKDYTYVEQYERTEGRHHLFL 225

Q9GTJ7 Q9GTJ7 PRELIMINARY; PRT; 230 AA.  
AC Q9GTJ7;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Red fluorescent protein.  
GN Name=FP593;  
OS Discosoma sp. SSAL-2000.  
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;  
OC Discosomatidae; Discosoma.  
OX NCBI\_TaxID=137428;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20434599; PubMed=10981720; DOI=10.1016/S0014-5793(00)01895-0;  
RA Fradkov A.F., Chen Y., Ding L., Barsova E.V., Matz M.V.,  
RA Lukyanov S.A.;  
RT "Novel fluorescent protein from Discosoma coral and its mutants  
possesses a unique far-red fluorescence.";  
RL FEBS Lett. 479:127-130(2000).  
DR EMBL; AF272711; AAG16224.1; -.  
DR HSSP; Q9U6Y8; 1GGX.  
DR GO; GO:0006091; P:energy pathways; IEA.  
DR InterPro; IPR009017; GFP\_like.  
DR InterPro; IPR011584; GFP\_related.  
DR InterPro; IPR000786; Green\_fl\_protein.  
DR Pfam; PF01353; GFP; 1.  
DR PRINTS; PR01229; GFLUORESCENT.  
SQ SEQUENCE 230 AA; 26370 MW; 5215B1B436D67E51 CRC64;  
  
Query Match 89.4%; Score 1085.5; DB 2; Length 230;  
Best Local Similarity 88.1%; Pred. No. 5.4e-87;  
Matches 199; Conservative 17; Mismatches 9; Indels 1; Gaps 1;  
  
QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVTKLVTKGGPLPFAWDI 60  
Db 1 MSCSKNVKEFMRFKVRMEGTVNGHEFEIKGEGRPYEGHCSVKLMVTKGGPLPFAFDI 60  
  
QY 61 LSPQFQYGSKVYVHPADIPDYKKLSFPEGFKWERMVNFEDGGVTVTQDSSLQDGCFIY 120  
Db 61 LSPQFQYGSKVYVHPADIPDYKKLSFPEGFKWERMVNFEDGGVTVTQDSSLKDGCFIY 120  
  
QY 121 KVKEFIGNVFSDGVPVQKKTMGWEASTERLYPRDGLKGEIHKALKKLDGGHYLVEFKSI 180  
Db 121 EVKFIGNVFSDGVPVQRRTRGWEASSERLYPRDGLKGDIIHMLRLEGGGHYLVEFKSI 180  
  
QY 181 YMAKKP-VQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMVKKPSVQLPGYIYVDSKLDMTSHNEDYTVVEQYEKTQGRHHPFI 226  
  
RESULT 6  
Q66PV1 PRELIMINARY; PRT; 221 AA.  
AC Q66PV1;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Chromoprotein.  
OS Acropora tenuis (Purple tipped acropora).  
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;  
OC Astrocoeniina; Acroporidae; Acropora.  
OX NCBI\_TaxID=70783;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Alieva N.O., Meleshkevitch E.A., Field S.F., Matz M.V.;  
RT "Survey of coral GFP-like proteins.";  
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY646074; AAU06853.1; -.  
DR InterPro; IPR009017; GFP\_like.  
DR InterPro; IPR011584; GFP\_related.  
DR Pfam; PF01353; GFP; 1.

DR PRINTS; PR01229; GFLUORESCENT.  
DR ProDom; PD013756; Green\_fl\_protein; 1.  
SQ SEQUENCE 221 AA; 25103 MW; 3BEA8B848C8F5DA4 CRC64;  
  
Query Match 66.2%; Score 804; DB 2; Length 221;  
Best Local Similarity 67.9%; Pred. No. 2.4e-62;  
Matches 148; Conservative 25; Mismatches 45; Indels 0; Gaps 0;  
  
QY 6 NVIKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVTKLVTKGGPLPFAWDILSPQF 65  
Db 2 SVIAKQMTYKVYMSGTVNGHYFEVEGDGKPKPYEGEQTVKLTVTTKGGPLPFAWDILSPQF 61  
  
QY 66 QYGSKVYVHPADIPDYKKLSFPEGFKWERMVNFEDGGVTVTQDSSLQDGCFIYKVKFI 125  
Db 62 QYGSIPFTKYPEDIPDYVKQSFPEGYTWERIMNFEDDAVCTVSNDDSIQGNCFIYHVKFS 121  
  
QY 126 GVNFPDGPVQKKTMGWEASTERLYPRDGLKGEIHKALKKLDGGHYLVEFKSIYMAKK 185  
Db 122 GVNFPDGPVQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGGHYLCEFEKSTYKAKK 181  
  
QY 186 PVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHL 223  
Db 182 PVRMPGYHYVDRKLDVTNHNKDYTSVEQCEISIARKPL 219  
  
RESULT 7  
Q66PU8 PRELIMINARY; PRT; 221 AA.  
AC Q66PU8;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Chromoprotein.  
OS Acropora aculeus.  
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;  
OC Astrocoeniina; Acroporidae; Acropora.  
OX NCBI\_TaxID=287157;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Alieva N.O., Meleshkevitch E.A., Field S.F., Matz M.V.;  
RT "Survey of coral GFP-like proteins.";  
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY646077; AAU06856.1; -.  
DR InterPro; IPR009017; GFP\_like.  
DR InterPro; IPR011584; GFP\_related.  
DR InterPro; IPR000786; Green\_fl\_protein.  
DR Pfam; PF01353; GFP; 1.  
DR PRINTS; PR01229; GFLUORESCENT.  
DR ProDom; PD013756; Green\_fl\_protein; 1.  
SQ SEQUENCE 221 AA; 24976 MW; 307DE4FC4B018B43 CRC64;  
  
Query Match 65.7%; Score 798; DB 2; Length 221;  
Best Local Similarity 69.0%; Pred. No. 8e-62;  
Matches 145; Conservative 24; Mismatches 41; Indels 0; Gaps 0;  
  
QY 6 NVIKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVTKLVTKGGPLPFAWDILSPQF 65  
Db 2 SVIAKQMTYKVYMSGTVNGHYFEVEGDGKPKPYEGEQTVKLTVTTKGGPLPFAWDILSPQS 61  
  
QY 66 QYGSKVYVHPADIPDYKKLSFPEGFKWERMVNFEDGGVTVTQDSSLQDGCFIYKVKFI 125  
Db 62 QYGSIPFTKYPDIPDYVKQSFPEGYTWERIMNFEDGAVCTVSNDDSIQGNCFIYNVKFS 121  
  
QY 126 GVNFPDGPVQKKTMGWEASTERLYPRDGLKGEIHKALKKLDGGHYLVEFKSIYMAKK 185  
Db 122 GLNFPDGPVQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGGHYLCEFEKSTYKAKK 181  
  
QY 186 PVQLPGYIYVDSKLDITSHNEDYTIIVEQYE 215  
Db 182 PVRMPGYHYVDRKLDVTNHNRDYTSVEQCE 211  
  
RESULT 8





AC P83690;  
DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Luminescent protein.  
OS Montipora efflorescens (Coral).  
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;  
OC Astrocoeniina; Acroporidae; Montipora.  
OX NCBI\_TaxID=105610;  
RN [1]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RX PubMed=12623015; DOI=10.1016/S0969-2126(03)00028-5;  
RA Prescott M., Ling M., Beddoe T., Oakley A.J., Dove S.,  
RA Hoegh-Guldberg O., Devenish R.J., Rossjohn J.;  
RT "The 2.2 a crystal structure of a pocilloporin pigment reveals a  
RT nonplanar chromophore conformation.";  
RL Structure 11:275-284(2003).  
DR PDB; lMOU; X-ray; A=-.  
DR PDB; lMOV; X-ray; A=1-219.  
DR GO; GO:0006091; P:energy pathways; IEA.  
DR InterPro; IPR009017; GFP\_like.  
DR InterPro; IPR011584; GFP\_related.  
DR InterPro; IPR000786; Green\_fl\_protein.  
DR Pfam; PF01353; GFP; 1.  
DR ProDom; PD013756; Green\_fl\_protein; 1.  
KW 3D-structure.  
SQ SEQUENCE 219 AA; 24569 MW; 434B2C95CB2DC7F9 CRC64;  
  
Query Match 61.9%; Score 751; DB 2; Length 219;  
Best Local Similarity 67.0%; Pred. No. 1e-57;  
Matches 140; Conservative 22; Mismatches 45; Indels 2; Gaps 1;  
  
QY 7 VIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPRYEGHNTVKKLVTKGGPLPFAWDILSPQFQ 66  
Db 3 VIATQMTYKVYMSGTVNGHYFEVEGDGKGRPYEGEQTVKLTVTKGGPLPFAWDILSP--Q 60  
  
QY 67 YGSKVYVVKHPADIPDYKKLSFPEGFKWERVMNFEDEGGVTVTQDSSLQDGCFFIYKVKFIG 126  
Db 61 CXSIPFTKYPEDIPDYVKQSFPEGFTWERIMNFEDEGAVCTVSNDSIIQNCFTYHVKFSG 120  
  
QY 127 VNFPSDGPVMQKTMGWEASTERLYPRDGLVKGKGEIHKALKLDGGHYLVVEFKSIYMAKKP 186  
Db 121 LNFPNGPVMQKTKQWEPHSERLFAFGGMLIGNFMALKLEGGGHYLCFEKTTYKAKKP 180  
  
QY 187 VOLPGYYVVDKLDITSHNEDYTYVEQYE 215  
Db 181 VKMPGYHYVDRKLDVTNHNKDYTSVEQCE 209  
  
RESULT 12  
Q66ND6 PRELIMINARY; PRT; 227 AA.  
AC Q66ND6;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Cyan fluorescent protein C1.  
OS Discosoma striata.  
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;  
OC Discosomatidae; Discosoma.  
OX NCBI\_TaxID=105400;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Carter R.W., Gibbs P.D.L., Schmale M.C.;  
RT "Cloning of Cnidarian Fluorescent Protein Genes.";  
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY679108; AAU04445.1; -.  
DR InterPro; IPR009017; GFP\_like.  
DR InterPro; IPR011584; GFP\_related.  
DR InterPro; IPR000786; Green\_fl\_protein.  
DR InterPro; IPR000169; Pept\_cys\_acsite.  
DR Pfam; PF01353; GFP; 1.  
DR PRINTS; PR01229; GFLUORESCENT.

DR ProDom; PD013756; Green\_fl\_protein; 1.  
DR PROSITE; PS00639; THIOI\_PROTEASE\_HIS; UNKNOWN\_1.  
SQ SEQUENCE 227 AA; 25640 MW; 568639306F44B244 CRC64;  
  
Query Match 61.4%; Score 745; DB 2; Length 227;  
Best Local Similarity 60.8%; Pred. No. 3.6e-57;  
Matches 135; Conservative 33; Mismatches 54; Indels 0; Gaps 0;  
  
QY 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPRYEGHNTVKKLVTKGGPLPFAWDI 60  
Db 1 MSWSKSVIKEEMLINLHLEGTFGNGHYFEIKGKGKGNPNEGTNTVLEVTKGGLPFGWHI 60  
  
QY 61 LSPQFYGSKVYVVKHPADIPDYKKLSFPEGFKWERVMNFEDEGGVTVTQDSSLQDGCFFIY 120  
Db 61 LCPQFYGKNKAFVHHPPDDIPDYKLSPFGYTWERSMHFEDGGLCCITNDISLTGNCIFY 120  
  
QY 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLVKGKGEIHKALKLDGGHYLVVEFKSI 180  
Db 121 DIKFTGLNFPNGPVMQKTTGWEPSTERLYPRDGLVIGDIHHALTVAGGGHYVCDIKTV 180  
  
QY 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTYVEQYERTEGRHH 222  
Db 181 YRAKXPVKMPGYHYVDPKLVIRSNDKFMKVEEHEIAVARXH 222  
  
RESULT 13  
Q9U6Y7 PRELIMINARY; PRT; 232 AA.  
AC Q9U6Y7;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Fluorescent protein FP483.  
OS Discosoma striata.  
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;  
OC Discosomatidae; Discosoma.  
OX NCBI\_TaxID=105400;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99436614; PubMed=10504696;  
RA Matz M.V., Fradkov A.F., Labas Y.A., Savitsky A.P., Zaratisky A.G.,  
RA Markelov M.L., Lukyanov S.A.;  
RT "Fluorescent proteins from nonbioluminescent Anthozoa species.";  
RL Nat. Biotechnol. 17:969-973(1999).  
DR EMBL; AF168420; AAF03370.1; -.  
DR HSSP; Q9U6Y8; 1GGX.  
DR GO; GO:0006091; P:energy pathways; IEA.  
DR InterPro; IPR009017; GFP\_like.  
DR InterPro; IPR011584; GFP\_related.  
DR InterPro; IPR000786; Green\_fl\_protein.  
DR Pfam; PF01353; GFP; 1.  
DR PRINTS; PR01229; GFLUORESCENT.  
DR ProDom; PD013756; Green\_fl\_protein; 1.  
SQ SEQUENCE 232 AA; 26435 MW; AA8F18EEE283CE4D CRC64;  
  
Query Match 60.1%; Score 729.5; DB 2; Length 232;  
Best Local Similarity 59.1%; Pred. No. 8.5e-56;  
Matches 133; Conservative 38; Mismatches 53; Indels 1; Gaps 1;  
  
QY 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPRYEGHNTVKKLVTKGGPLPFAWDI 60  
Db 1 MSCSKSVIKEEMLIDLHLEGTFGNGHYFEIKGKGKQPNEGTNTVLEVTKGGLPFGWHI 60  
  
QY 61 LSPQFYGSKVYVVKHPADIPDYKKLSFPEGFKWERVMNFEDEGGVTVTQDSSLQDGCFFIY 120  
Db 61 LCPQFYGKNKAFVHHPPDDIPDYKLSPFGYTWERSMHFEDGGLCCITNDISLTGNCIFY 120  
  
QY 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLVKGKGEIHKALKLDGGHYLVVEFKSI 180  
Db 121 DIKFTGLNFPNGPVMQKTTGWEPSTERLYPRDGLVIGDIHHALTVAGGGHYACDIKT 180  
  
QY 181 YMAKK-PVOLPGYYVVDKLDITSHNEDYTYVEQYERTEGRHFLF 224

Db 181 YRAKKAALKMFGYHYVDTKLVIWNNDKEFMKVEEHEIAVARHHPF 225

RESULT 14

Q963F5 PRELIMINARY; PRT; 225 AA.

AC Q963F5;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Green fluorescent protein.

OS Montastraea cavernosa (great star coral).

OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;

OC Faviina; Faviidae; Montastraea.

OX NCBI\_TaxID=63558;

RN [1]

RP SEQUENCE FROM N.A.

RA Lesser M.P., Barry T.M., Mazel C., Matz M.V., Lukyanov S.A.,

RA Falkowski P., Gorbunov M., Kolber Z.;

RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF384683; AAK62982.2; -.

DR HSSP; Q9U6Y8; 1GGX.

DR GO; GO:0006091; P:energy pathways; IEA.

DR InterPro; IPR009017; GFP\_like.

DR InterPro; IPR011584; GFP\_related.

DR InterPro; IPR000786; Green\_fl\_protein.

DR Pfam; PF01353; GFP; 1.

DR PRINTS; PR01229; GFLUORESCENT.

SQ SEQUENCE 225 AA; 25847 MW; 77DE7D7C616929AF CRC64;

Query Match 56.8%; Score 690; DB 2; Length 225;

Best Local Similarity 55.0%; Pred. No. 2.4e-52;

Matches 120; Conservative 44; Mismatches 54; Indels 0; Gaps 0;

QY 6 NVIKEFMRFKVRMEGTVNGHEFEIEGEGRPPYEGHNTVKLVTKGGPLPFAWDILSPQF 65

Db 2 SVIKPIMEIKLRMQGVVNGHKFKVKGEGEGKPFEGTQTINLTVKEGAPLPFAYDILTSAF 61

QY 66 QYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVTVTQDSSLQDGCFFIYKVKFI 125

Db 62 QYGNRVFTKYPPDDIPDYFKQTFPEGYSWERIMAYEDQSICTATSDIKMEGDCFIYEIQFH 121

QY 126 GVNFPDGPVMQKTMGWEASTERLYPRDGVLKGEIHKALKKDGGHYLVFEFKSIYMAKK 185

Db 122 GVNFPNGPVMQKTLKWEPTKMYVRDGVLGDNMALLLEGGGHYRCDFRSTYKAKK 181

QY 186 PVQLPGYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHL 223

Db 182 RVQLPDYHFVDHRIEILSHDNDYNTVKLSEDAEARYSM 219

RESULT 15

Q66ND3 PRELIMINARY; PRT; 225 AA.

AC Q66ND3;

DT 25-OCT-2004 (TrEMBLrel. 28, Created)

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Green fluorescent protein G2.

OS Montastraea cavernosa (great star coral).

OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;

OC Faviina; Faviidae; Montastraea.

OX NCBI\_TaxID=63558;

RN [1]

RP SEQUENCE FROM N.A.

RA Carter R.W., Gibbs P.D.L., Schmale M.C.;

RT "Cloning of Cnidarian Fluorescent Protein Genes.";

RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY679111; AAU04448.1; -.

DR InterPro; IPR009017; GFP\_like.

DR InterPro; IPR011584; GFP\_related.

DR InterPro; IPR000786; Green\_fl\_protein.

DR Pfam; PF01353; GFP; 1.

DR PRINTS; PR01229; GFLUORESCENT.

DR ProDom; PD013756; Green\_fl\_protein; 1.

SQ SEQUENCE 225 AA; 25910 MW; A744E8D6A6C423AE CRC64;

Query Match 56.3%; Score 684; DB 2; Length 225;

Best Local Similarity 55.0%; Pred. No. 7.9e-52;

Matches 120; Conservative 41; Mismatches 57; Indels 0; Gaps 0;

QY 6 NVIKEFMRFKVRMEGTVNGHEFEIEGEGRPPYEGHNTVKLVTKGGPLPFAWDILSPQF 65

Db 2 SVIKPDMKIKLRMEGAVNGHNFIEGEGKGPFEQTQTINLTVKEGGPLPFAYDILTAAF 61

QY 66 QYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVTVTQDSSLQDGCFFIYKVKFI 125

Db 62 QYGNRAFTKYPRDIADYFKQSFPEGYSWERSMTYEDQGICICIKSDIRMEGDCFIYEIRYD 121

QY 126 GVNFPDGPVMQKTMGWEASTERLYPRDGVLKGEIHKALKKDGGHYLVFEFKSIYMAKK 185

Db 122 GVNFPSPGPVMQKTLKWEPTKMYVRDGVLKGEVNMALLLEGGGHYRCDFRSTYKAKK 181

QY 186 PVQLPGYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHL 223

Db 182 RVQLPDYHFVDHRIEILSHDNDYNTVKLSENAEARYSM 219

Search completed: June 30, 2005, 15:49:35

Job time : 176 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 15:31:23 ; Search time 162 Seconds  
(without alignments)  
537.167 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVKEFRFKVRMEG.....EDYTIQEYERTEGRHHLFL 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1214	100.0	225	3	AAY99836 Discosoma
2	1214	100.0	225	3	AAB01622 Discosoma
3	1214	100.0	225	4	AAG65509 Anthozoan
4	1214	100.0	225	5	ABB08834 Yeast opt
5	1214	100.0	225	5	AAE28833 Discosoma
6	1214	100.0	225	5	AAE17540 Discosoma
7	1214	100.0	225	5	AAO18270 Discosoma
8	1214	100.0	225	6	AAE34962 Discosoma
9	1214	100.0	225	7	ADC24126 Discosoma
10	1214	100.0	225	7	ABW00918 Discosoma
11	1214	100.0	225	7	ADF70403 Discosoma
12	1214	100.0	225	7	ADH34489 Discosoma
13	1214	100.0	225	7	ADL46203 Discosoma
14	1214	100.0	225	7	ADN33979 Wild-type
15	1214	100.0	225	8	ADI36421 Discosoma
16	1214	100.0	225	8	ADM97769 D sp red
17	1214	100.0	225	8	ADQ28780 Jellyfish
18	1214	100.0	487	5	ABB08821 Autofluor
19	1214	100.0	506	5	ABB08822 Autofluor
20	1214	100.0	547	5	ABB08823 Autofluor
21	1211	99.8	225	5	AAE28920 Discosoma
22	1211	99.8	225	7	ABW00937 Discosoma
23	1211	99.8	225	7	ABW00929 Discosoma
24	1211	99.8	225	7	ABW00938 Discosoma
25	1211	99.8	225	7	ADH34498 Discosoma

26	1211	99.8	225	7	ADH34499	Adh34499 Discosoma
27	1210	99.7	225	5	AAE28919	Aae28919 Discosoma
28	1210	99.7	225	7	ABW00930	Abw00930 Discosoma
29	1210	99.7	226	4	AAG65510	Aag65510 Anthozoan
30	1210	99.7	242	7	ADE24109	Ade24109 Discosoma
31	1210	99.7	545	7	ADL18132	Adl18132 RFP:PS(NI
32	1210	99.7	548	7	ADL18156	Adl18156 RFP:PS(HI
33	1209	99.6	225	5	AAE28922	Aae28922 Discosoma
34	1209	99.6	225	7	ABW00936	Abw00936 Discosoma
35	1209	99.6	225	7	ABW00931	Abw00931 Discosoma
36	1209	99.6	226	5	ABB08835	Abb08835 Yeast opt
37	1208	99.5	225	7	ABW00932	Abw00932 Discosoma
38	1208	99.5	225	7	ABW00935	Abw00935 Discosoma
39	1208	99.5	225	7	ABW00939	Abw00939 Discosoma
40	1208	99.5	225	7	ADH34500	Adh34500 Discosoma
41	1207	99.4	225	5	AAE28921	Aae28921 Discosoma
42	1207	99.4	225	5	AAE17541	Aae17541 Discosoma
43	1207	99.4	225	7	ABW00941	Abw00941 Discosoma
44	1207	99.4	225	7	ABW00940	Abw00940 Discosoma
45	1207	99.4	225	7	ADL46222	Adl46222 Discosoma

ALIGNMENTS

RESULT 1  
AAY99836  
ID AAY99836 standard; protein; 225 AA.  
XX AC AAY99836;  
XX  
DT 12-SEP-2003 (revised)  
DT 19-SEP-2000 (first entry)  
XX  
DE Discosoma sp. "red" novel fluorescent protein drFP583.  
XX  
KW Anthozoa; drFP583; fluorescent protein; non-bioluminescent organism;  
KW Fluorescent labeling.  
XX  
OS Discosoma sp; "red".  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 122 /note= "encoded by TC"  
FT Misc-difference 127 /note= "encoded by GTTG"  
FT  
XX WO200034326-A1.  
XX  
PD 15-JUN-2000.  
XX  
PF 10-DEC-1999; 99WO-US029473.  
XX  
PR 11-DEC-1998; 98US-00210330.  
PR 14-OCT-1999; 99US-00418529.  
XX  
PA (CLON-) CLONTECH LAB INC.  
XX  
PI Lukyanov SA, Fradkov AF, Labas YA, Matz MV, Green G, Chen Y;  
PI Ding L;  
XX  
DR WPI; 2000-423381/36.  
XX  
PT Novel fluorescent protein from non-bioluminescent Discosoma sp. red,  
PT useful for fluorescent labeling and as markers.  
XX  
PS Claim 20; Page 74-75; 86pp; English.  
XX  
CC The present sequence is a novel fluorescent protein (nFP) encoded by the  
CC full-length cDNA drFP583. drFP583 was isolated from Discosoma sp. "red",  
CC a non-bioluminescent species of the Class Anthozoa. Fluorescent proteins  
CC can be used in fluorescent labeling, a useful tool for marking a protein,  
CC cell or organism of interest. Unlike other markers used in protein

CC labeling, such as beta-galactosidase and luciferase, fluorescent proteins  
CC do not require an exogenous cofactor or substrate. Methods involving  
CC fluorescent proteins are also less laborious and less difficult to  
CC control than the traditional methods of fluorescent labeling, where a  
CC protein of interest is purified and then covalently conjugated to a  
CC fluorophore derivative. Novel fluorescent proteins isolated from species  
CC of the Class Anthozoa can be used as markers for gene expression and  
CC protein localization studies, and in fluorescence resonance energy  
CC transfer (FRET) reactions. They may have improved properties and better  
CC suitability for larger excitations compared to prior art fluorescent  
CC proteins such as green fluorescent protein. (Updated on 12-SEP-2003 to  
CC standardise OS field)

XX  
SQ Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 3; Length 225;  
Best Local Similarity 100.0%; Pred. No. 8.9e-128;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTGGPLPFAWDI 60  
Db 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTGGPLPFAWDI 60  
QY 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERMVNFEDGGVTVTQDSSLQDGCFIY 120  
Db 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERMVNFEDGGVTVTQDSSLQDGCFIY 120  
QY 121 KVKEFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLDGGHYLVEFKSI 180  
Db 121 KVKEFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 225

RESULT 2  
AAB01622  
ID AAB01622 standard; protein; 225 AA.  
XX AAB01622;  
XX  
DT 12-DEC-2000 (first entry)  
XX  
DE Discosoma sp. red fluorescent protein drFP583.  
XX  
KW Anthozoa; fluorescent protein; fluorescence labeling; amFP486; cFP484;  
KW zFP506; zFP538; drFP583; dsFP483; asFP600; dgFP512; dmFP592.  
XX  
OS Discosoma sp.  
XX  
PN WO200034526-A1.  
XX  
PD 15-JUN-2000.  
XX  
PF 10-DEC-1999; 99WO-US029405.  
XX  
PR 11-DEC-1998; 98US-00210330.  
XX  
PA (CLON-) CLONTECH LAB INC.  
XX  
PI Lukyanoy SA, Fradkov AF, Labas YA, Matz MV;  
XX  
DR WPI; 2000-423451/36.  
XX  
PT Novel method for identifying a DNA sequence encoding fluorescent proteins  
PT from non-bioluminescent Anthozoa which are useful for fluorescent  
PT labeling and as markers.  
XX  
PS Claim 3; Page 68-69; 73pp; English.  
XX  
CC The present sequence is Discosoma sp. red fluorescent protein drFP583. It  
CC was isolated using the Aequoria victoria green fluorescent protein (GFP)

CC sequence, which was used to design PCR primers which might isolate other  
CC fluorescent proteins from a number of species of Anthozoa. These were  
CC Anemonia majano, Clavularia sp., Zoanthus sp., Discosoma sp. and Anemonia  
CC sulcata. The cDNA obtained was then screened in the search for sequences  
CC encoding fluorescent proteins. The other proteins found in this manner  
CC were cFP484, zFP506, zFP538, amFP486, dsFP483, asFP600, dgFP512 and  
CC dmFP592. These proteins can be used as fluorescent labels (for gene  
CC expression and protein localisation studies and in fluorescence resonance  
CC energy transfer (FRET) studies) in place of fluorophore derivatives and  
CC luciferases, as these involve laborious processes and the latter require  
CC cofactors. They can also be used in place of GFP, which is too stable to  
CC be useful when studying short-term or repetitive events

XX  
SQ Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 3; Length 225;  
Best Local Similarity 100.0%; Pred. No. 8.9e-128;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTGGPLPFAWDI 60  
Db 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTGGPLPFAWDI 60  
QY 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERMVNFEDGGVTVTQDSSLQDGCFIY 120  
Db 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERMVNFEDGGVTVTQDSSLQDGCFIY 120  
QY 121 KVKEFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLDGGHYLVEFKSI 180  
Db 121 KVKEFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 225

RESULT 3  
AAG65509  
ID AAG65509 standard; protein; 225 AA.  
XX AAG65509;  
XX  
DT 30-NOV-2001 (first entry)  
XX  
DE Anthozoa red fluorescent protein sequence.  
XX  
KW Fluorescent protein; Anthozoa; fluorescence; marker; FRET; red.  
XX  
OS Anthozoa.  
XX  
PN WO200162919-A1.  
XX  
PD 30-AUG-2001.  
XX  
PF 13-FEB-2001; 2001WO-US004625.  
XX  
PR 23-FEB-2000; 2000US-0184732P.  
XX  
PA (AURO-) AURORA BIOSCIENCES CORP.  
XX  
PI Nelson D, Zamaira E, Tsien R;  
XX  
DR WPI; 2001-557704/62.  
XX  
PT Proteins for Fluorescence Resonance Energy Transfer (FRET) comprise  
PT functional red fluorescent proteins, and the encoding nucleic acids, with  
PT key mutations for improving the proteins function.  
XX  
PS Disclosure; Page 85; 90pp; English.  
XX  
CC The invention provides a nucleic acid encoding functional red fluorescent  
CC protein (II) that differs from the sequence of an Anthozoan red  
CC fluorescent protein by at least one amino acid substitution, and with

CC different fluorescent properties. The red fluorescent protein of the  
CC invention can be expressed by standard recombinant methodology. (II) are  
CC used a fluorescent markers and FRET partners. It is used for identifying  
CC protein-protein interactions. (II) is also suitable for multiplexed  
CC fluorescent analysis and FRET-based applications using existing Aequorea  
CC fluorescent proteins. (II) has improved brightness, reduced spectral  
CC cross talk, and is rapidly and efficiently expressed in mammalian cells.  
CC The key mutations in the encoding nucleic acids provide improved folding,  
CC brightness, and create (II) with sharper, more defined excitation and  
CC emission peaks when expressed in mammalian cells. The present sequence  
CC represents an anthozoan fluorescent protein

XX  
SQ Sequence 225 AA;  
  
Query Match 100.0%; Score 1214; DB 4; Length 225;  
Best Local Similarity 100.0%; Pred. No. 8.9e-128;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVCLKVTKGGPLPFAWDI 60  
Db 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVCLKVTKGGPLPFAWDI 60  
  
QY 61 LSPQFQYGSKVYVKHPADIPDYKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGCIFY 120  
Db 61 LSPQFQYGSKVYVKHPADIPDYKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGCIFY 120  
  
QY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKDKDGGHYLVEFKSI 180  
Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKDKDGGHYLVEFKSI 180  
  
QY 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 4  
ABB08834  
ID ABB08834 standard; protein; 225 AA.  
XX  
AC ABB08834;  
XX  
DT 29-MAY-2002 (first entry)  
XX  
DE Yeast optimised RFP SEQ ID NO 17.  
XX  
KW Yeast; red fluorescent protein; RFP; plant; transgenic; GFP;  
KW Saccharomyces cerevisiae; Nicotiana tabacum; Arabidopsis thaliana;  
KW Escherichia coli; green fluorescent protein; biotechnology.  
XX  
OS Anthozoa.  
XX  
PN DE20001395-U1.  
XX  
PD 15-MAR-2001.  
XX  
PF 27-JAN-2000; 2000DE-02001395.  
XX  
PR 27-JAN-2000; 2000DE-02001395.  
XX  
PA (GPCB-) GPC BIOTECH AG.  
XX  
DR WPI; 2002-228394/29.  
DR N-PSDB; ABA95905, ABA95921, ABA95922.  
XX  
PT New DNA encoding red fluorescent protein, useful as marker in  
PT biotechnology, has sequence optimized for expression in eukaryotes,  
PT especially yeast or plants.  
XX  
PS Disclosure; Page 13-14; 19pp; German.  
XX  
CC The invention relates to DNA (I) containing either sequence ABA95905 or  
CC sequence ABA95906 encoding a yeast optimised Red Fluorescent Protein  
CC (yRFP). (I) are used to express red fluorescent protein (RFP) in

CC eukaryotes, especially yeast, especially Saccharomyces cerevisiae and  
CC plants, especially dicotyledonous plants including Nicotiana tabacum or  
CC Arabidopsis thaliana and also in prokaryotes, especially bacteria,  
CC especially Escherichia coli. RFP is useful in the same way as green  
CC fluorescent protein but is more generally applicable in modern  
CC biotechnology. (I) are optimised for expression in yeast and so generate  
CC RFP at higher levels with stronger fluorescence and thus lowers the  
CC detection limit and gives a better signal-to-noise ratio. The present  
CC sequence is that of the yeast optimised RFP

XX  
SQ Sequence 225 AA;  
  
Query Match 100.0%; Score 1214; DB 5; Length 225;  
Best Local Similarity 100.0%; Pred. No. 8.9e-128;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVCLKVTKGGPLPFAWDI 60  
Db 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVCLKVTKGGPLPFAWDI 60  
  
QY 61 LSPQFQYGSKVYVKHPADIPDYKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGCIFY 120  
Db 61 LSPQFQYGSKVYVKHPADIPDYKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGCIFY 120  
  
QY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKDKDGGHYLVEFKSI 180  
Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKDKDGGHYLVEFKSI 180  
  
QY 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 5  
AAE28833  
ID AAE28833 standard; protein; 225 AA.  
XX  
AC AAE28833;  
XX  
DT 27-DEC-2002 (first entry)  
XX  
DE Discosoma sp. drFP583 (NFP-6) wild-type protein.  
XX  
KW Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
KW fluorescence activated cell sorting application; fluorescent timer;  
KW biosensor; fluorescence resonance energy transfer application; FRET;  
KW colouring agent; recombinant DNA application; analyte detection assay;  
KW sunscreen; second messenger detector; drFP583 protein; NFP-6.  
XX  
OS Discosoma sp.  
XX  
PN WO200268459-A2.  
XX  
PD 06-SEP-2002.  
XX  
PF 20-FEB-2002; 2002WO-US005749.  
XX  
PR 21-FEB-2001; 2001US-0270983P.  
PR 04-DEC-2001; 2001US-00006922.  
XX  
PA (CLON-) CLONTECH LAB INC.  
XX  
PI Lukyanov S, Lukyanov K, Yanushevich Y, Savitsky A, Fradkov A;  
XX WPI; 2002-691654/74.  
DR N-PSDB; AAD46278.  
XX  
PT New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
PT of an aggregating Cnidarian chromo- or fluorescent protein or mutant for  
PT analyte detection assays or fluorescence activated cell sorting  
PT applications.  
XX  
PS Disclosure; Page 70-71; 80pp; English.



XX The invention relates to nucleic acid molecules encoding non-aggregating  
CC chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
CC useful in analyte detection assays, as colouring agents, as markers in  
CC recombinant DNA applications, as sunscreens or filters, in fluorescence  
CC resonance energy transfer (FRET) applications, as biosensors in  
CC prokaryotic and eukaryotic cells, in screening assays, as second  
CC messenger detectors, in fluorescence activated cell sorting applications,  
CC in protease cleavage assays or as fluorescent timers. The present  
CC sequence is Discosoma sp. drFP583 (NFP-6) wild-type protein of the  
CC invention  
XX  
SQ Sequence 225 AA;  
  
Query Match 100.0%; Score 1214; DB 5; Length 225;  
Best Local Similarity 100.0%; Pred. No. 8.9e-128;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVKTKGGPLPFAWDI 60  
Db 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVKTKGGPLPFAWDI 60  
  
QY 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFDGGVVTVTQDSSLQDGCFIY 120  
Db 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFDGGVVTVTQDSSLQDGCFIY 120  
  
QY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLDGGHYLVEFKSI 180  
Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLDGGHYLVEFKSI 180  
  
QY 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIQEYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIQEYERTEGRHHLFL 225  
  
RESULT 6  
AAE17540  
ID AAE17540 standard; protein; 225 AA.  
XX  
AC AAE17540;  
XX  
DT 22-APR-2002 (first entry)  
XX  
DE Discosoma sp. humanised wild-type Anthozoa protein drFP583.  
XX  
KW Fluorescent timer protein; protein movement; translocation; trafficking;  
KW promoter activity; gene expression; transgenic plant; gene modification;  
KW protein age; anthozoa protein; drFP583.  
XX  
OS Discosoma sp.  
XX  
PN WO200196373-A2.  
XX  
PD 20-DEC-2001.  
XX  
PF 13-JUN-2001; 2001WO-US019097.  
XX  
PR 14-JUN-2000; 2000US-0211607P.  
XX  
PA (CLON-) CLONTECH LAB INC.  
XX  
PI Pradkov AF, Terskikh A;  
XX  
DR WPI; 2002-154595/20.  
DR N-PSDB; AAD28207.  
XX  
PT New fluorescent timer proteins comprising an emission spectrum that  
PT changes over time from a first wavelength to a second wavelength, useful  
PT for monitoring intracellular protein movement, translocation, trafficking  
PT or stability.  
XX  
PS Example 1; Fig 1; 89pp; English.  
XX

CC The invention relates to a fluorescent timer protein having an emission  
CC spectrum that changes over time after synthesis from a first wavelength  
CC to a second wavelength. The fluorescent timer proteins are useful in  
CC monitoring the activity of a promoter, determining the age of a protein,  
CC identifying an agent that modulates the activity of a promoter and in  
CC enriching a population of cells comprising a fluorescent timer protein.  
CC The fluorescent timer proteins are also useful for assessing gene  
CC expression during development of a multicellular organism or during  
CC cellular differentiation, in response to a drug or other inducer of  
CC promoter activity, as a reporter to serve as a read-out of promoter  
CC activity, monitoring intracellular protein movement or translocation,  
CC protein trafficking, or protein stability, to investigate temporal  
CC aspects of the activity of a regulatory element, for determining cell  
CC fate during development and organ remodelling, in spatial and temporal  
CC visualisation of newly synthesised proteins and accumulated proteins, and  
CC in distinguishing between newly formed and pre-existing structures, e.g.  
CC membrane junctions and extracellular matrix components. The fluorescent  
CC timer proteins may further be used to investigate components where photobleaching  
CC techniques are employed, as detectable labels, as selectable markers, as  
CC biosensors in prokaryotic and eukaryotic cells, in protease cleavage  
CC assays, and as second messenger detectors. The nucleic acids can be used  
CC to generate transgenic, non-human plants or animals or site-specific gene  
CC modifications in cell lines. The present sequence is Discosoma sp.  
CC humanised wild-type Anthozoa protein drFP583 used for generating  
CC fluorescent proteins  
XX  
SQ Sequence 225 AA;  
  
Query Match 100.0%; Score 1214; DB 5; Length 225;  
Best Local Similarity 100.0%; Pred. No. 8.9e-128;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVKTKGGPLPFAWDI 60  
Db 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVKTKGGPLPFAWDI 60  
  
QY 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFDGGVVTVTQDSSLQDGCFIY 120  
Db 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFDGGVVTVTQDSSLQDGCFIY 120  
  
QY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLDGGHYLVEFKSI 180  
Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLDGGHYLVEFKSI 180  
  
QY 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIQEYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIQEYERTEGRHHLFL 225  
  
RESULT 7  
AAO18270  
ID AAO18270 standard; protein; 225 AA.  
XX  
AC AAO18270;  
XX  
DT 26-SEP-2002 (first entry)  
XX  
DE Discosoma red fluorescent protein.  
XX  
KW Yeast; RAD54; promoter; genotoxicity cassette; cytotoxicity cassette;  
KW modified yeast strain; environmental pollution.  
XX  
OS Discosoma sp.  
XX  
PN DE10061872-A1.  
XX  
PD 20-JUN-2002.  
XX  
PF 12-DEC-2000; 2000DE-01061872.  
XX  
PR 12-DEC-2000; 2000DE-01061872.  
XX  
PA (LICH/) LICHTENBERG-FRATE H.

XX Lichtenberg-Frate H;  
PI  
XX  
DR WPI; 2002-539633/58.  
DR N-PSDB; AAL47952.  
XX  
PT Modified yeast strain, useful for detecting toxic compounds in  
PT environment, contains integrated cassettes responsive to genotoxic and  
PT cytotoxic compounds.  
XX  
PS Disclosure; Page 21-22; 34pp; German.  
XX  
CC The present invention relates to a modified yeast strain that contains,  
CC integrated stably and functionally in its genome, a genotoxicity cassette  
CC and a cytotoxicity cassette, each comprising a promoter and reporter  
CC gene, both of which are different in the two cassettes. The modified  
CC yeast strain is used to detect environmental pollution, especially  
CC genotoxic and/or cytotoxic substances in complex environmental  
CC contaminants, especially organic compounds, but also (non-)ionising  
CC radiation and chemical carcinogens. Particular applications are in  
CC monitoring (waste) water (e.g. as an early warning system), medical  
CC toxicology screening and for industrial process control. The present  
CC sequence is a marker protein suitable for use in the cassettes of the  
CC present invention  
XX  
SQ Sequence 225 AA;  
Query Match 100.0%; Score 1214; DB 5; Length 225;  
Best Local Similarity 100.0%; Pred. No. 8.9e-128;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVTKGGPLPFAWDI 60  
Db 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVTKGGPLPFAWDI 60  
QY 61 LSPQFQYGSKVYVVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGCIFY 120  
Db 61 LSPQFQYGSKVYVVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGCIFY 120  
QY 121 KVKGFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLVKGKALVTKGGHYLVFEKSI 180  
Db 121 KVKGFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLVKGKALVTKGGHYLVFEKSI 180  
QY 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
RESULT 8  
AAE34962  
ID AAE34962 standard; protein; 225 AA.  
XX  
AC AAE34962;  
XX  
DT 28-MAY-2003 (first entry)  
XX  
DE Discosoma species red fluorescent protein (RFP).  
XX  
KW Phosphorylation indicator; fluorescent protein; detection; phosphatase;  
KW kinase; red fluorescent protein; RFP.  
XX  
OS Discosoma sp.  
XX  
PN WO200295058-A2.  
XX  
PD 28-NOV-2002.  
XX  
PF 24-MAY-2002; 2002WO-US016955.  
XX  
PR 24-MAY-2001; 2001US-00865291.  
XX  
PA (REGC ) UNIV CALIFORNIA.  
XX

PI Taien RY, Ting AY, Zhang J;  
XX  
DR WPI; 2003-148474/14.  
DR N-PSDB; AAD53432.  
XX  
PT Novel chimeric phosphorylation indicators, useful for detecting  
PT kinase/phosphatase in samples, has donor molecule, phosphorylatable  
PT domain, phosphoaminoacid binding domain, and acceptor molecule, in  
PT operative linkage.  
XX  
PS Disclosure; Col 65-66; 38pp; English.  
XX  
CC The present invention relates to chimeric phosphorylation indicators  
CC comprising a phosphorylation polypeptide and a fluorescent protein or in  
CC operative linkage, a donor molecule, a phosphorylatable domain, a  
CC phosphoaminoacid binding domain (PAABD) and an acceptor molecule. The  
CC phosphorylation indicators of the invention are useful for detecting  
CC kinases or phosphatases in a biological sample. They are also useful in  
CC high throughput analysis e.g. for detecting a kinase inhibitor or  
CC phosphatase inhibitor. The present sequence is Discosoma species red  
CC fluorescent protein (RFP) used in the invention  
XX  
SQ Sequence 225 AA;  
Query Match 100.0%; Score 1214; DB 6; Length 225;  
Best Local Similarity 100.0%; Pred. No. 8.9e-128;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVTKGGPLPFAWDI 60  
Db 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVTKGGPLPFAWDI 60  
QY 61 LSPQFQYGSKVYVVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGCIFY 120  
Db 61 LSPQFQYGSKVYVVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGCIFY 120  
QY 121 KVKGFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLVKGKALVTKGGHYLVFEKSI 180  
Db 121 KVKGFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLVKGKALVTKGGHYLVFEKSI 180  
QY 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
RESULT 9  
ADC24126  
ID ADC24126 standard; protein; 225 AA.  
XX  
AC ADC24126;  
XX  
DT 18-DEC-2003 (first entry)  
XX  
DE Discosoma wild-type red fluorescent protein.  
XX  
KW Discosoma red fluorescent protein; DsRed; AB interface; AC interface;  
KW fluorescent protein variant; transcription induction detection;  
KW fluorescence energy resonance transfer; FRET; protein kinase;  
KW protein phosphatase; ion indicator.  
XX  
OS Discosoma.  
XX  
PN US2003059835-A1.  
XX  
PD 27-MAR-2003.  
XX  
PF 10-APR-2002; 2002US-00121258.  
XX  
PR 26-FEB-2001; 2001US-00794308.  
PR 24-MAY-2001; 2001US-00866538.  
XX  
PA (TSIE/) TSIE R Y.  
PA (CAMP/) CAMPBELL R E.





KW cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;  
KW GFPuv; Enhanced GFP; EGFP.  
XX  
OS Discosoma sp.  
XX  
PN WO2003071272-A1.  
XX  
PD 28-AUG-2003.  
XX  
PF 21-FEB-2003; 2003WO-JP001901.  
XX  
PR 22-FEB-2002; 2002JP-00045728.  
PR 23-JUL-2002; 2002JP-00213949.  
PR 11-OCT-2002; 2002JP-00298237.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Hinuma S, Fujii R, Ogi K, Komatsu H, Kawamata Y, Hosoya M;  
XX  
DR WPI; 2003-697654/66.  
DR N-PSDB; ADF70404.  
XX  
PT Transformation of cells with a fusion protein of an orphan receptor  
PT protein with a fluorescent protein useful for identification of ligands  
PT to the orphan receptor.  
XX  
PS Disclosure; SEQ ID NO 26; 594pp; Japanese.  
XX  
CC This invention relates to a novel method of identifying ligands to an  
CC orphan receptor protein which comprises transforming cells with DNA  
CC encoding a fusion protein of the orphan receptor with a fluorescent  
CC protein, so that the fusion protein is expressed in the cells (or cell  
CC membranes isolated from them) and contacting the cells with the potential  
CC ligand to be tested. A suitable fluorescent protein for incorporation in  
CC the fusion protein is green fluorescent protein (GFP), for example GFP-1,  
CC wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the  
CC identification of ligands binding to an orphan receptor protein.  
XX  
SQ Sequence 225 AA;  
  
Query Match 100.0%; Score 1214; DB 7; Length 225;  
Best Local Similarity 100.0%; Pred. No. 8.9e-128;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGEGRPEYEGHNTVKLVKTKGGPLPFAWDI 60  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGEGRPEYEGHNTVKLVKTKGGPLPFAWDI 60  
  
QY 61 LSPQFQYGSKVYVXHPADIPDYKLSFPEGFKWERMVNFEDGGVTVTQDSSLQDGCFIY 120  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
61 LSPQFQYGSKVYVXHPADIPDYKLSFPEGFKWERMVNFEDGGVTVTQDSSLQDGCFIY 120  
  
QY 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
  
QY 181 YMAKKPVQLPGYYVVDKSLDITSHNEDYIVEQYERTEGRHHFL 225  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
181 YMAKKPVQLPGYYVVDKSLDITSHNEDYIVEQYERTEGRHHFL 225  
  
RESULT 12  
ADH34489  
ID ADH34489 standard; protein; 225 AA.  
XX  
AC ADH34489;  
XX  
DT 11-MAR-2004 (first entry)  
XX  
DE Discosoma sp. red fluorescent protein DsRed (wild-type).  
XX  
KW Chromoprotein; fluorescent protein; CP; FP; interconverted mutant;  
KW Cnidarian; Anthozoan; labelling; colouring agents; pigment;

KW analyte detection assay; selectable marker; sunscreen; selective filter;  
KW fluorescence resonance energy transfer; FRET; biosensor;  
KW whole cell marker; second messenger detector; in vivo marker;  
KW fluorescence activated cell sorting; fluorescent timer;  
KW red fluorescent protein; DsRed.  
XX  
OS Discosoma sp.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 146 /note= "Key position for chromoprotein/fluorescent  
FT protein activity. Corresponds to GFP residue 148"  
FT Misc-difference 161 /note= "Key position for chromoprotein/fluorescent  
FT protein activity. Corresponds to GFP residue 165"  
FT Misc-difference 163 /note= "Key position for chromoprotein/fluorescent  
FT protein activity. Corresponds to GFP residue 167"  
FT Misc-difference 197 /note= "Key position for chromoprotein/fluorescent  
FT protein activity. Corresponds to GFP residue 203"  
XX  
PN WO2003057833-A2.  
XX  
PD 17-JUL-2003.  
XX  
PF 23-DEC-2002; 2002WO-US041418.  
XX  
PR 26-DEC-2001; 2001US-0343128P.  
PA (CLON-) CLONTECH LAB INC.  
XX  
XX Bulina ME, Chudakov D, Lukyanov KA;  
PI WPI; 2003-607998/57.  
XX  
DR Novel nucleic acid encoding interconverted mutant of chromo-or  
XX fluorescent protein which are useful as biosensors, coloring agents.  
PT  
PT  
XX  
PS Example 1; Fig 1; 56pp; English.  
XX  
CC The invention relates to interconverted mutants of chromoproteins (CP) or  
CC fluorescent proteins (FP) and nucleic acids encoding them. The mutant is  
CC derived from a Cnidarian species, preferably a non-bioluminescent  
CC Cnidarian species, and most preferably an Anthozoan species. The  
CC invention is based on the finding that although green fluorescent protein  
CC (GFP)-like chromoproteins and fluorescent proteins exhibit some degree of  
CC homology, there are certain positions (referred to as 148, 165, 167 and  
CC 203; numbering corresponds to GFP) that are occupied by noticeably  
CC different residues in the two types of proteins. Mutagenesis of the  
CC residues in these key positions in, for example, a fluorescent protein,  
CC to those found in a chromoprotein is therefore proposed to confer  
CC chromoprotein activity on the fluorescent protein mutant, with  
CC chromoproteins being able to be converted into fluorescent proteins in a  
CC similar manner. The invention also relates to expression constructs,  
CC vectors, host cells and host cell progeny comprising a nucleic acid of  
CC the invention; the recombinant production of an interconverted  
CC chromoprotein or fluorescent protein mutant; and antibodies specific for  
CC interconverted mutant proteins of the invention. The interconverted  
CC mutants are useful in any application that employs a chromoprotein or  
CC fluorescent protein. Fluorescent protein mutants having chromoprotein  
CC activity can be useful as colouring agents in, for example, food  
CC compositions, pharmaceuticals, cosmetics and living organisms. Proteins  
CC with chromoprotein activity are also useful as labels in biological  
CC analyte detection assays, as selectable markers in recombinant DNA  
CC applications (e.g. the production of transgenic cells and organisms), and  
CC are also useful as sunscreens and selective filters. Chromoprotein  
CC mutants having fluorescent protein activity useful in fluorescence  
CC resonance energy transfer (FRET) applications, as biosensors in  
CC prokaryotic and eukaryotic cells, as markers of whole cells to detect  
CC changes in multicellular reorganisation and migration, as second  
CC messenger detectors, as in vivo markers in animals (e.g., transgenic  
CC animals), in fluorescence activated cell sorting applications, in

CC protease cleavage assays, and in assays to determine the phospholipid  
CC composition in biological membranes. Proteins with fluorescent protein  
CC activity can also be used as fluorescent timers, where the switch of one  
CC fluorescent colour to another (e.g., green to red) is concomitant with  
CC the ageing of the protein and is useful for determination of the  
CC activation or deactivation of gene expression. The present sequence  
CC represents a wild-type red fluorescent protein, DsRed, from *Discosoma* sp.  
CC that was used as a parent sequence for the generation of mutant proteins  
CC in an example of the invention.  
XX  
SQ Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 7; Length 225;  
Best Local Similarity 100.0%; Pred. No. 8.9e-128;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLKVTGGPLPFAWDI 60  
Db 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLKVTGGPLPFAWDI 60  
Qy 61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGCIFY 120  
Db 61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGCIFY 120  
Qy 121 KVKEFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLDGGHYLVEFKSI 180  
Db 121 KVKEFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLDGGHYLVEFKSI 180  
Qy 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 225

RESULT 13  
ADL46203  
ID ADL46203 standard; protein; 225 AA.  
XX  
AC ADL46203;  
XX  
DT 20-MAY-2004 (first entry)  
XX  
DE *Discosoma* red fluorescent (DsRed) protein.  
XX  
KW red fluorescent protein; DsRed; fluorescence; red wavelength;  
KW oligomerization; tetramerization; immunoassay; hybridization assay.  
XX  
OS *Discosoma* sp.  
XX  
PN WO2003086446-A1.  
XX  
PD 23-OCT-2003.  
XX  
PF 09-APR-2003; 2003WO-US010879.  
XX  
PR 10-APR-2002; 2002US-00121258.  
PR 29-JUL-2002; 2002US-00209208.  
XX  
PA (REGC ) UNIV CALIFORNIA.  
XX  
XX  
PI Tsien RY, Campbell RE, Baird GS;  
XX  
DR WPI; 2003-845265/78.  
DR N-PSDB; ADL46204.  
XX

PT New monomeric and dimeric Anthozoan fluorescent protein variants with  
PT reduced propensity to oligomerize, and encoding polynucleotides, useful  
PT in molecular biology, e.g. in immunoassays or in tracking protein  
PT movement in cells.  
XX  
PS Claim 1; SEQ ID NO 1; 166pp; English.  
XX  
CC The invention relates to a polynucleotide sequence encoding a *Discosoma*  
CC red fluorescent protein (DsRed) variant having a reduced propensity to

CC oligomerize. The protein variant comprises one or more amino acid  
CC substitutions at the AB and/or AC interface(s) of the wild-type DsRed  
CC sequence, where the substitutions result in reduced propensity of the  
CC DsRed variant to form tetramers and where the variant displays detectable  
CC fluorescence of at least one red wavelength. The composition and methods  
CC are useful in producing red fluorescent proteins having reduced  
CC propensity for oligomerization, especially tetramerization. The protein  
CC may be used in molecular biology and in other scientific applications,  
CC such as in immunoassays or hybridization assays, or in tracking the  
CC movement of proteins in cells. This sequence corresponds to the DsRed  
CC protein.  
XX  
SQ Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 7; Length 225;  
Best Local Similarity 100.0%; Pred. No. 8.9e-128;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLKVTGGPLPFAWDI 60  
Db 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLKVTGGPLPFAWDI 60  
Qy 61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGCIFY 120  
Db 61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGCIFY 120  
Qy 121 KVKEFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLDGGHYLVEFKSI 180  
Db 121 KVKEFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLDGGHYLVEFKSI 180  
Qy 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 225

RESULT 14  
ADN33979  
ID ADN33979 standard; protein; 225 AA.  
XX  
AC ADN33979;  
XX  
DT 01-JUL-2004 (first entry)  
XX  
DE Wild-type DsRed protein.  
XX  
KW Cnidarian; fluorescence resonance energy transfer; FRET; wild-type DsRed.  
XX  
OS *Discosoma* sp.  
XX  
PN WO2003054158-A2.  
XX  
PD 03-JUL-2003.  
XX  
PF 18-DEC-2002; 2002WO-US040539.  
XX  
PR 19-DEC-2001; 2001US-0341723P.  
XX  
PA (UYCH-) UNIV CHICAGO.  
XX  
PI Bevis B, Glick B;  
XX  
DR WPI; 2003-569236/53.  
DR N-PSDB; ADN33978.  
XX

PT Novel nucleic acid encoding a rapidly maturing chromo- or fluorescent  
PT mutant of a Cnidarian chromo- or fluorescent protein or its mutant,  
PT useful for applications involving chromo- or fluorescent proteins.  
XX  
PS Claim 8; SEQ ID NO 2; 65pp; English.  
XX  
CC The present invention relates to nucleic acid that encodes a rapidly  
CC maturing chromo or fluorescent mutant of a Cnidarian chromo- or  
CC fluorescent protein or its mutant. The protein is useful in applications

CC involving nucleic acid encoding a chromo- or fluorescent protein and is  
CC useful for producing a chromo and/or fluorescent protein which involves  
CC growing the cell, whereby the protein is expressed, and isolating the  
CC protein substantially free of other proteins. The protein is useful in  
CC applications involving chromo- or fluorescent protein and is useful as  
CC PCR primers, hybridization probes, etc. The expression cassettes are  
CC useful for synthesizing related proteins. The chromoproteins are useful  
CC as coloring agents which are capable of imparting color or pigment to a  
CC particular composition of matter e.g. food compositions, pharmaceuticals,  
CC cosmetics, living organisms, e.g., animals and plants. The chromoproteins  
CC may also find use as labels in analyte detection assays, e.g. assays for  
CC biological analytes of interest and as selectable markers in recombinant  
CC DNA applications, e.g. the production of transgenic cells and organisms.  
CC The fluorescent proteins find use in a variety of different applications,  
CC e.g. in fluorescence resonance energy transfer (FRET) applications, as  
CC biosensors in prokaryotic and eukaryotic cells, in applications involving  
CC the automated screening of arrays of cells expressing fluorescent  
CC reporting groups by using microscopic imaging and electronic analysis, as  
CC second messenger detectors, and in fluorescence activated cell sorting  
CC applications and as in vivo marker in animals. The fluorescent proteins  
CC also find use in protease cleavage assays. The proteins can also be used  
CC is assays to determine the phospholipid composition in biological  
CC membranes and as a fluorescent timer. The present sequence represents the  
CC wild-type DSRED.

XX  
SQ Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 7; Length 225;  
Best Local Similarity 100.0%; Pred. No. 8.9e-128;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVTKGGPLPFAWDI 60  
Db 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVTKGGPLPFAWDI 60  
Qy 61 LSPQFQYGSKVYVVKHPADIPDYKKLSFPEGFKWERMVNFEDGGVVTVTQDSSLQDGCIFY 120  
Db 61 LSPQFQYGSKVYVVKHPADIPDYKKLSFPEGFKWERMVNFEDGGVVTVTQDSSLQDGCIFY 120  
Qy 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKDKGGHYLVEFKSI 180  
Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKDKGGHYLVEFKSI 180  
Qy 181 YMAKKPVQLPGYYVVDVKLDTITSHNEDYTIQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYYVVDVKLDTITSHNEDYTIQYERTEGRHHLFL 225

RESULT 15

ADI36421  
ID ADI36421 standard; protein; 225 AA.

AC ADI36421;

XX  
DT 22-APR-2004 (first entry)

XX  
DE Discosoma sp. red fluorescent protein (RED).

XX  
KW Fluorobody; binding ligand; green fluorescent protein; GFP;  
KW target detection; red fluorescent protein; RED.

OS  
XX Discosoma sp.

XX  
FH Key Location/Qualifiers  
FT Region 22. .26  
FT /note= "Loop region"  
FT Region 37. .39  
FT /note= "Loop region"  
FT Region 75. .81  
FT /note= "Loop region"  
FT Region 100. .103  
FT /note= "Loop region"  
FT Region 114. .117

FT Region /note= "Loop region"  
FT 153. .156  
FT /note= "Loop region"  
FT 167. .170  
FT /note= "Loop region"  
FT 185. .192  
FT /note= "Loop region"  
FT 204. .209  
FT /note= "Loop region"

XX US2003203355-A1.

PN 30-OCT-2003.

XX 24-APR-2002; 2002US-00132067.

XX 24-APR-2002; 2002US-00132067.

XX (LALA-) LOS ALAMOS NAT LAB.  
PA (REGC ) UNIV CALIFORNIA.

XX Bradbury AM, Zeytun A, Waldo GS;

XX WPI; 2004-154325/15.

DR N-PSDB; ADI36420.

XX  
PT Novel binding ligand with intrinsic fluorescence and comprising  
PT fluorescent protein having heterologous binding sites, useful for  
PT detecting target molecule.

XX  
PS Claim 2; SEQ ID NO 4; 23pp; English.

XX  
CC The invention relates to binding ligands (fluorobodies) with intrinsic  
CC fluorescence, which comprises green fluorescent protein (GFP) having  
CC heterologous binding sites. The binding ligand is useful for detecting  
CC the target molecule and is efficiently detects the target molecule. The  
CC present sequence is Discosoma sp. red fluorescent protein (RED) used in  
CC the exemplification of the invention.

XX  
SQ Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 8; Length 225;  
Best Local Similarity 100.0%; Pred. No. 8.9e-128;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVTKGGPLPFAWDI 60  
Db 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVTKGGPLPFAWDI 60

Qy 61 LSPQFQYGSKVYVVKHPADIPDYKKLSFPEGFKWERMVNFEDGGVVTVTQDSSLQDGCIFY 120  
Db 61 LSPQFQYGSKVYVVKHPADIPDYKKLSFPEGFKWERMVNFEDGGVVTVTQDSSLQDGCIFY 120

Qy 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKDKGGHYLVEFKSI 180  
Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKDKGGHYLVEFKSI 180

Qy 181 YMAKKPVQLPGYYVVDVKLDTITSHNEDYTIQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYYVVDVKLDTITSHNEDYTIQYERTEGRHHLFL 225

Search completed: June 30, 2005, 15:46:34  
Job time : 165 secs



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 15:43:54 ; Search time 161 Seconds  
(without alignments)  
539.155 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMRFKVRMEG.....EDYTIVEQYERTEGRHHLFL 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1722976 seqs, 385795295 residues

Total number of hits satisfying chosen parameters: 1722976

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA:\*

1:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17:	/cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18:	/cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
19:	/cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20:	/cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1214	100.0	225	9	US-09-999-745-67
2	1214	100.0	225	10	US-09-866-538-12
3	1214	100.0	225	10	US-09-794-308-12
4	1214	100.0	225	10	US-09-865-291-12
5	1214	100.0	225	13	US-10-006-922-12
6	1214	100.0	225	14	US-10-081-864-8
7	1214	100.0	225	14	US-10-121-258-1
8	1214	100.0	225	14	US-10-315-920-2
9	1214	100.0	225	15	US-10-132-067-4
10	1214	100.0	225	15	US-10-370-570-56
11	1214	100.0	225	15	US-10-406-618-32

12	1214	100.0	225	16	US-10-433-640-13	Sequence 13, Appl
13	1214	100.0	225	16	US-10-311-030-7	Sequence 7, Appli
14	1214	100.0	225	16	US-10-845-484-3	Sequence 3, Appli
15	1214	100.0	225	16	US-10-885-988-12	Sequence 12, Appl
16	1214	100.0	225	17	US-10-656-029-22	Sequence 22, Appl
17	1214	100.0	225	17	US-10-857-622-12	Sequence 12, Appl
18	1214	100.0	225	17	US-10-505-486-26	Sequence 26, Appl
19	1214	100.0	487	15	US-10-343-977-1	Sequence 1, Appli
20	1214	100.0	487	17	US-10-343-977-1	Sequence 1, Appli
21	1214	100.0	506	15	US-10-343-977-2	Sequence 2, Appli
22	1214	100.0	506	17	US-10-343-977-2	Sequence 2, Appli
23	1214	100.0	547	15	US-10-343-977-3	Sequence 3, Appli
24	1214	100.0	547	17	US-10-343-977-3	Sequence 3, Appli
25	1210	99.7	226	16	US-10-311-030-9	Sequence 9, Appli
26	1210	99.7	240	14	US-10-152-296-2	Sequence 2, Appli
27	1210	99.7	240	16	US-10-739-656-2	Sequence 2, Appli
28	1210	99.7	545	14	US-10-214-932-52	Sequence 52, Appl
29	1210	99.7	548	14	US-10-214-932-76	Sequence 76, Appl
30	1207	99.4	225	14	US-10-121-258-20	Sequence 20, Appl
31	1207	99.4	225	14	US-10-315-920-4	Sequence 4, Appli
32	1196	98.5	236	16	US-10-314-936-2	Sequence 2, Appli
33	1196	98.5	236	16	US-10-314-936-4	Sequence 4, Appli
34	1196	98.5	236	19	US-11-021-014-2	Sequence 2, Appli
35	1196	98.5	236	19	US-11-021-014-4	Sequence 4, Appli
36	1195.5	98.5	226	16	US-10-724-178-12	Sequence 12, Appl
37	1191	98.1	225	14	US-10-315-920-6	Sequence 6, Appli
38	1186	97.7	225	15	US-10-442-148A-7	Sequence 7, Appli
39	1186	97.7	239	15	US-10-442-148A-8	Sequence 8, Appli
40	1160	95.6	225	14	US-10-121-258-4	Sequence 4, Appli
41	1153	95.0	225	14	US-10-121-258-24	Sequence 24, Appl
42	1132	93.2	225	15	US-10-423-688A-41	Sequence 41, Appl
43	1121	92.3	226	14	US-10-121-258-6	Sequence 6, Appli
44	1119	92.2	225	13	US-10-006-922-44	Sequence 44, Appl
45	1119	92.2	225	14	US-10-081-864-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1  
US-09-999-745-67  
; Sequence 67, Application US/09999745  
; Patent No. US20020157120A1  
; GENERAL INFORMATION:  
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: Tsien, Roger Y.  
; APPLICANT: Baird, Geoffrey  
; TITLE OF INVENTION: CIRCULARLY PERMUTED FLUORESCENT PROTEIN INDICATORS  
; FILE REFERENCE: REGEN1470-1  
; CURRENT APPLICATION NUMBER: US/09/999,745  
; CURRENT FILING DATE: 2001-10-23  
; PRIOR APPLICATION NUMBER: 09/316,920  
; PRIOR FILING DATE: 1999-05-21  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 67  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Discosoma sp.  
US-09-999-745-67

Query Match 100.0%; Score 1214; DB 9; Length 225;  
Best Local Similarity 100.0%; Pred. No. 6.1e-114;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGERPYEGHNTVKLVTKGGPLPFAWDI	60
Db	1	MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGERPYEGHNTVKLVTKGGPLPFAWDI	60
QY	61	LSPQFQYGSKVYVKKHPADIPDYKLSFPEGFKWERMNFEDGGVVTVTQDSSLQDGCIFY	120
Db	61	LSPQFQYGSKVYVKKHPADIPDYKLSFPEGFKWERMNFEDGGVVTVTQDSSLQDGCIFY	120

QY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKDKDGGHYLVEFKSI 180  
|||||  
Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKDKDGGHYLVEFKSI 180  
|||||  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 225  
|||||  
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 225  
|||||  
RESULT 2  
US-09-866-538-12  
; Sequence 12, Application US/09866538  
; Publication No. US20030032088A1  
; GENERAL INFORMATION:  
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: TSUEN, Roger  
; APPLICANT: Campbell, Robert  
; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS  
; FILE REFERENCE: REGEN1530-2  
; CURRENT APPLICATION NUMBER: US/09/866,538  
; CURRENT FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 12  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Discosoma sp.  
US-09-866-538-12

Query Match 100.0%; Score 1214; DB 10; Length 225;  
Best Local Similarity 100.0%; Pred. No. 6.1e-114;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVTKGGPLPFAWDI 60  
|||||  
Db 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVTKGGPLPFAWDI 60  
|||||  
QY 61 LSPQFQYGSKVYVVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVTVTQDSSLQDGCIFY 120  
|||||  
Db 61 LSPQFQYGSKVYVVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVTVTQDSSLQDGCIFY 120  
|||||  
QY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKDKDGGHYLVEFKSI 180  
|||||  
Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKDKDGGHYLVEFKSI 180  
|||||  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 225  
|||||  
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 225  
|||||

RESULT 3  
US-09-794-308-12  
; Sequence 12, Application US/09794308  
; Publication No. US20030170911A1  
; GENERAL INFORMATION:  
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: TSUEN, Roger  
; APPLICANT: ZACHARIAS, David  
; APPLICANT: BAIRD, Geoffrey  
; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS  
; FILE REFERENCE: REGEN1530  
; CURRENT APPLICATION NUMBER: US/09/794,308  
; CURRENT FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 12  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Discosoma sp.  
US-09-794-308-12

Query Match 100.0%; Score 1214; DB 10; Length 225;  
Best Local Similarity 100.0%; Pred. No. 6.1e-114;

Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVTKGGPLPFAWDI 60  
|||||  
Db 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVTKGGPLPFAWDI 60  
|||||  
QY 61 LSPQFQYGSKVYVVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVTVTQDSSLQDGCIFY 120  
|||||  
Db 61 LSPQFQYGSKVYVVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVTVTQDSSLQDGCIFY 120  
|||||  
QY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKDKDGGHYLVEFKSI 180  
|||||  
Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKDKDGGHYLVEFKSI 180  
|||||  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 225  
|||||  
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 225  
|||||

RESULT 4  
US-09-865-291-12  
; Sequence 12, Application US/09865291  
; Publication No. US20030186229A1  
; GENERAL INFORMATION:  
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: TSUEN, Roger  
; APPLICANT: TING, Alice  
; APPLICANT: ZHANG, Jin  
; TITLE OF INVENTION: EMISSION RATIONETRIC INDICATORS OF PHOSPHORYLATION  
; FILE REFERENCE: REGEN1550  
; CURRENT APPLICATION NUMBER: US/09/865,291  
; CURRENT FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 12  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Discosoma sp.  
US-09-865-291-12

Query Match 100.0%; Score 1214; DB 10; Length 225;  
Best Local Similarity 100.0%; Pred. No. 6.1e-114;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVTKGGPLPFAWDI 60  
|||||  
Db 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVTKGGPLPFAWDI 60  
|||||  
QY 61 LSPQFQYGSKVYVVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVTVTQDSSLQDGCIFY 120  
|||||  
Db 61 LSPQFQYGSKVYVVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVTVTQDSSLQDGCIFY 120  
|||||  
QY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKDKDGGHYLVEFKSI 180  
|||||  
Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKDKDGGHYLVEFKSI 180  
|||||  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 225  
|||||  
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 225  
|||||

RESULT 5  
US-10-006-922-12  
; Sequence 12, Application US/10006922  
; Publication No. US20020197676A1  
; GENERAL INFORMATION:  
; APPLICANT: Lukyanov, Sergey A  
; APPLICANT: Fradkov, Arcady F.  
; APPLICANT: Labas, Yulii A.  
; APPLICANT: Matz, Mikhail V.  
; APPLICANT: Tersikh, Alexey  
; TITLE OF INVENTION: No. US20020197676A1el Chromophores/Fluorophores and  
; TITLE OF INVENTION: Methods for Using the Same

```

; FILE REFERENCE: CLON-035CIP
; CURRENT APPLICATION NUMBER: US/10/006,922
; CURRENT FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 09/120,330
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/457,898
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,144
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,477
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/457,556
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/444,338
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma species
US-10-006-922-12

Query Match      100.0%; Score 1214; DB 13; Length 225;
Best Local Similarity 100.0%; Pred. No. 6.1e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVCLKVTKGGPLPFAWDI 60
Db      1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVCLKVTKGGPLPFAWDI 60

QY      61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVTVTTQDSSLQDGCIFY 120
Db      61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVTVTTQDSSLQDGCIFY 120

QY      121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLDGGHYLVEFKSI 180
Db      121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLDGGHYLVEFKSI 180

QY      181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
Db      181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 6
US-10-081-864-8
; Sequence 8, Application US/10081864
; Publication No. US20030022287A1
; GENERAL INFORMATION:
; APPLICANT: Lukyanov, Sergey
; APPLICANT: Lukyanov, Konstantin
; APPLICANT: Yanushevich, Yuriy
; APPLICANT: Savitsky, Alexandr
; APPLICANT: Fradkov, Arcady
; TITLE OF INVENTION: No. US20030022287A1 Aggregating Fluorescent Proteins and
; FILE REFERENCE: CLON-067
; CURRENT APPLICATION NUMBER: US/10/081,864
; CURRENT FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: 10/006,922
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/270,983
; PRIOR FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
US-10-081-864-8

Query Match      100.0%; Score 1214; DB 14; Length 225;
Best Local Similarity 100.0%; Pred. No. 6.1e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVCLKVTKGGPLPFAWDI 60
Db      1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVCLKVTKGGPLPFAWDI 60

QY      61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVTVTTQDSSLQDGCIFY 120
Db      61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVTVTTQDSSLQDGCIFY 120

QY      121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLDGGHYLVEFKSI 180
Db      121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLDGGHYLVEFKSI 180

QY      181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
Db      181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

US-10-081-864-8
```

```

Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVCLKVTKGGPLPFAWDI 60
Db      1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVCLKVTKGGPLPFAWDI 60

QY      61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVTVTTQDSSLQDGCIFY 120
Db      61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVTVTTQDSSLQDGCIFY 120

QY      121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLDGGHYLVEFKSI 180
Db      121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLDGGHYLVEFKSI 180

QY      181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
Db      181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 7
US-10-121-258-1
; Sequence 1, Application US/10121258
; Publication No. US20030059835A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; FILE REFERENCE: UC083.1CP2CP1
; CURRENT APPLICATION NUMBER: US/10/121,258
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(225)
; OTHER INFORMATION: wild-type DsRed
US-10-121-258-1

Query Match      100.0%; Score 1214; DB 14; Length 225;
Best Local Similarity 100.0%; Pred. No. 6.1e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVCLKVTKGGPLPFAWDI 60
Db      1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVCLKVTKGGPLPFAWDI 60

QY      61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVTVTTQDSSLQDGCIFY 120
Db      61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVTVTTQDSSLQDGCIFY 120

QY      121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLDGGHYLVEFKSI 180
Db      121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLDGGHYLVEFKSI 180

QY      181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
Db      181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

US-10-315-920-2
; Sequence 2, Application US/10315920
; Publication No. US20030175809A1
; GENERAL INFORMATION:
```



```

; APPLICANT: Fradkov, Arcady Fedorovich
; APPLICANT: Terskikh, Alexey
; TITLE OF INVENTION: FLUORESCENT TIMER PROTEINS AND METHODS
; TITLE OF INVENTION: FOR THEIR USE
; FILE REFERENCE: CLON-077CIP
; CURRENT APPLICATION NUMBER: US/10/315,920
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/211,607
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: PCT/US01/19097
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
US-10-315-920-2

Query Match      100.0%; Score 1214; DB 14; Length 225;
Best Local Similarity 100.0%; Pred. No. 6.1e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVCLKVTGGPLPFAWDI 60
   |||||||
Db 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVCLKVTGGPLPFAWDI 60

QY 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERMVNFEDGGVTVTQDSSLQDGCFIY 120
   |||||||
Db 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERMVNFEDGGVTVTQDSSLQDGCFIY 120

QY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLDGGHYLVEFKSI 180
   |||||||
Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLDGGHYLVEFKSI 180

QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
   |||||||
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 9
US-10-132-067-4
; Sequence 4, Application US/10132067
; Publication No. US20030203355A1
; GENERAL INFORMATION:
; APPLICANT: Bradbury, Andrew
; APPLICANT: Zeytun, Ahmet
; APPLICANT: Waldo, Geoffrey
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Fluorobodies: Binding Ligands With Intrinsic
; TITLE OF INVENTION: Fluorescence
; FILE REFERENCE: 021362-000600US
; CURRENT APPLICATION NUMBER: US/10/132,067
; CURRENT FILING DATE: 2002-04-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 4
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
; FEATURE:
; OTHER INFORMATION: red fluorescent protein (dsRED)
US-10-132-067-4

Query Match      100.0%; Score 1214; DB 15; Length 225;
Best Local Similarity 100.0%; Pred. No. 6.1e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVCLKVTGGPLPFAWDI 60
   |||||||
Db 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVCLKVTGGPLPFAWDI 60

QY 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERMVNFEDGGVTVTQDSSLQDGCFIY 120
   |||||||
Db 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERMVNFEDGGVTVTQDSSLQDGCFIY 120

QY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLDGGHYLVEFKSI 180
   |||||||
Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLDGGHYLVEFKSI 180

QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
   |||||||
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
```

```

Db 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERMVNFEDGGVTVTQDSSLQDGCFIY 120
   |||||||
QY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLDGGHYLVEFKSI 180
   |||||||
Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLDGGHYLVEFKSI 180
   |||||||
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
   |||||||
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
   |||||||

RESULT 10
US-10-370-570-56
; Sequence 56, Application US/10370570
; Publication No. US20030219717A1
; GENERAL INFORMATION:
; APPLICANT: DAHL, Soren Weis et al.
; TITLE OF INVENTION: FLUOROPHORE COMPLEMENTATION PRODUCTS
; FILE REFERENCE: 3759-0130P
; CURRENT APPLICATION NUMBER: US/10/370,570
; CURRENT FILING DATE: 2003-02-24
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 56
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
US-10-370-570-56

Query Match      100.0%; Score 1214; DB 15; Length 225;
Best Local Similarity 100.0%; Pred. No. 6.1e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVCLKVTGGPLPFAWDI 60
   |||||||
Db 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVCLKVTGGPLPFAWDI 60
   |||||||
QY 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERMVNFEDGGVTVTQDSSLQDGCFIY 120
   |||||||
Db 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERMVNFEDGGVTVTQDSSLQDGCFIY 120
   |||||||
QY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLDGGHYLVEFKSI 180
   |||||||
Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLDGGHYLVEFKSI 180
   |||||||
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
   |||||||
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
   |||||||

RESULT 11
US-10-406-618-32
; Sequence 32, Application US/10406618
; Publication No. US20030219814A1
; GENERAL INFORMATION:
; APPLICANT: Wan, David Chi-Cheong
; APPLICANT: Ip, Denis Tsz-Ming
; APPLICANT: The Chinese University of Hong Kong
; TITLE OF INVENTION: No. US20030219814A1el Fluorescent Proteins
; FILE REFERENCE: 016285-34-1
; CURRENT APPLICATION NUMBER: US/10/406,618
; CURRENT FILING DATE: 2003-04-02
; PRIOR APPLICATION NUMBER: US 60/387,968
; PRIOR FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: US 60/370,598
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial Sequence
```

```

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Spontaneously
; OTHER INFORMATION: Fluorescent protein DsRed.
US-10-406-618-32

Query Match      100.0%; Score 1214; DB 15; Length 225;
Best Local Similarity 100.0%; Pred. No. 6.1e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVKTKGGPLPFAWDI 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVKTKGGPLPFAWDI 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERMNFEDGGVVTVTQDSSLQDGCIFY 120
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERMNFEDGGVVTVTQDSSLQDGCIFY 120
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGVLGKGEIHKALKLDGGHYLVEFKSI 180
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGVLGKGEIHKALKLDGGHYLVEFKSI 180
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 12
US-10-433-640-13
; Sequence 13, Application US/10433640
; Publication No. US20040115792A1
; GENERAL INFORMATION:
; APPLICANT: Lichtenberg-Frate, Hella
; TITLE OF INVENTION: YEAST STRAIN FOR TESTING THE GENOTOXICITY AND CYTOTOXICITY OF
; TITLE OF INVENTION: COMPLEX ENVIRONMENTAL CONTAMINATIONS
; FILE REFERENCE: 1487/3
; CURRENT APPLICATION NUMBER: US/10/433,640
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/14610
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: DE 10061872.3
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
US-10-433-640-13

Query Match      100.0%; Score 1214; DB 16; Length 225;
Best Local Similarity 100.0%; Pred. No. 6.1e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVKTKGGPLPFAWDI 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVKTKGGPLPFAWDI 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERMNFEDGGVVTVTQDSSLQDGCIFY 120
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERMNFEDGGVVTVTQDSSLQDGCIFY 120
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGVLGKGEIHKALKLDGGHYLVEFKSI 180
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGVLGKGEIHKALKLDGGHYLVEFKSI 180
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 13
US-10-311-030-7
; Sequence 7, Application US/10311030
```

```

; Publication No. US20040171107A1
; GENERAL INFORMATION:
; APPLICANT: Nelson, David
; APPLICANT: Zambara, Elize
; APPLICANT: Tsien, Roger
; TITLE OF INVENTION: MODIFIED FLUORESCENT PROTEINS
; FILE REFERENCE: 15916-032US1
; CURRENT APPLICATION NUMBER: US/10/311,030
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: PCT/US01/04625
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: US 60/184,732
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-311-030-7

Query Match      100.0%; Score 1214; DB 16; Length 225;
Best Local Similarity 100.0%; Pred. No. 6.1e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVKTKGGPLPFAWDI 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVKTKGGPLPFAWDI 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERMNFEDGGVVTVTQDSSLQDGCIFY 120
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERMNFEDGGVVTVTQDSSLQDGCIFY 120
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGVLGKGEIHKALKLDGGHYLVEFKSI 180
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGVLGKGEIHKALKLDGGHYLVEFKSI 180
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 14
US-10-845-484-3
; Sequence 3, Application US/10845484
; Publication No. US20040248180A1
; GENERAL INFORMATION:
; APPLICANT: Bulina, Maria
; APPLICANT: Chudakov, Dmitry
; APPLICANT: Lukyanov, Konstantin
; TITLE OF INVENTION: Mutant Chromophores/Fluorophores and
; TITLE OF INVENTION: Methods for Making and Using the Same
; FILE REFERENCE: CLON 092
; CURRENT APPLICATION NUMBER: US/10/845,484
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: 60/343128
; PRIOR FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: PCT/US02/41418
; PRIOR FILING DATE: 2002-12-23
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
US-10-845-484-3

Query Match      100.0%; Score 1214; DB 16; Length 225;
Best Local Similarity 100.0%; Pred. No. 6.1e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVTKGGPLPFAWDI 60  
Db 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVTKGGPLPFAWDI 60  
QY 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGCFIY 120  
Db 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGCFIY 120  
QY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVEFKSI 180  
Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 15  
US-10-885-988-12  
; Sequence 12, Application US/10885988  
; Publication No. US20040259165A1  
; GENERAL INFORMATION:  
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: TSUEN, Roger  
; APPLICANT: Campbell, Robert  
; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS  
; FILE REFERENCE: REGEN1530-2  
; CURRENT APPLICATION NUMBER: US/10/885,988  
; CURRENT FILING DATE: 2004-07-06  
; PRIOR APPLICATION NUMBER: US/09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 12  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Discosoma sp.  
US-10-885-988-12

Query Match 100.0%; Score 1214; DB 16; Length 225;  
Best Local Similarity 100.0%; Pred. No. 6.1e-114;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVTKGGPLPFAWDI 60  
Db 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVTKGGPLPFAWDI 60  
QY 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGCFIY 120  
Db 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGCFIY 120  
QY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVEFKSI 180  
Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2005, 15:41:04 ; Search time 43 Seconds  
(without alignments)  
390.606 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMRFKVRMEG.....EDYTIQEYERTEGRHHLFL 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1210	99.7	240	4 US-10-152-296-2	Sequence 2, Appli
2	485.5	40.0	238	3 US-09-277-716-16	Sequence 16, Appl
3	485.5	40.0	238	4 US-09-609-161B-16	Sequence 16, Appl
4	485.5	40.0	238	4 US-09-626-581D-65	Sequence 65, Appl
5	485.5	40.0	238	4 US-09-415-765B-65	Sequence 65, Appl
6	485.5	40.0	238	4 US-09-626-580C-65	Sequence 65, Appl
7	473.5	39.0	224	4 US-09-977-897-3	Sequence 3, Appli
8	473.5	39.0	224	4 US-09-977-897-13	Sequence 13, Appl
9	473.5	39.0	225	4 US-09-977-897-14	Sequence 14, Appl
10	473.5	39.0	226	4 US-09-977-897-15	Sequence 15, Appl
11	473.5	39.0	227	4 US-09-977-897-16	Sequence 16, Appl
12	473.5	39.0	228	4 US-09-977-897-17	Sequence 17, Appl
13	473.5	39.0	229	4 US-09-977-897-18	Sequence 18, Appl
14	473.5	39.0	230	4 US-09-977-897-19	Sequence 19, Appl
15	473.5	39.0	231	4 US-09-977-897-20	Sequence 20, Appl
16	473.5	39.0	232	4 US-09-977-897-21	Sequence 21, Appl
17	473.5	39.0	233	4 US-09-977-897-22	Sequence 22, Appl
18	473.5	39.0	234	4 US-09-977-897-9	Sequence 9, Appli
19	473.5	39.0	234	4 US-09-977-897-23	Sequence 23, Appl
20	473.5	39.0	235	4 US-09-977-897-8	Sequence 8, Appli
21	473.5	39.0	235	4 US-09-977-897-24	Sequence 24, Appl
22	473.5	39.0	236	4 US-09-977-897-7	Sequence 7, Appli
23	473.5	39.0	236	4 US-09-977-897-25	Sequence 25, Appl
24	473.5	39.0	237	4 US-09-977-897-6	Sequence 6, Appli
25	473.5	39.0	237	4 US-09-977-897-26	Sequence 26, Appl
26	473.5	39.0	238	3 US-09-277-716-32	Sequence 32, Appl
27	473.5	39.0	238	4 US-09-609-161B-32	Sequence 32, Appl

28	473.5	39.0	238	4 US-09-839-650-3	Sequence 3, Appli
29	473.5	39.0	238	4 US-09-977-897-5	Sequence 5, Appli
30	473.5	39.0	238	4 US-09-977-897-27	Sequence 27, Appl
31	473.5	39.0	239	4 US-09-977-897-2	Sequence 2, Appli
32	473	39.0	219	4 US-09-977-897-4	Sequence 4, Appli
33	473	39.0	231	4 US-09-977-897-12	Sequence 12, Appl
34	473	39.0	232	4 US-09-977-897-11	Sequence 11, Appl
35	473	39.0	233	4 US-09-977-897-10	Sequence 10, Appl
36	215.5	17.8	238	4 US-09-023-946B-23	Sequence 23, Appl
37	214.5	17.7	238	4 US-09-023-946B-28	Sequence 28, Appl
38	213.5	17.6	240	4 US-09-129-192C-49	Sequence 49, Appl
39	212.5	17.5	238	2 US-08-818-604-32	Sequence 32, Appl
40	212.5	17.5	238	3 US-08-819-612-22	Sequence 22, Appl
41	212.5	17.5	238	4 US-09-316-919-3	Sequence 3, Appli
42	212.5	17.5	238	4 US-09-346-946-32	Sequence 32, Appl
43	212.5	17.5	238	4 US-09-023-946B-22	Sequence 22, Appl
44	212.5	17.5	238	4 US-09-316-920A-3	Sequence 3, Appli
45	212.5	17.5	238	4 US-09-872-364-22	Sequence 22, Appl

ALIGNMENTS

RESULT 1  
US-10-152-296-2  
; Sequence 2, Application US/101522296  
; Patent No. 6723537  
; GENERAL INFORMATION:  
; APPLICANT: Peelle, Beau  
; APPLICANT: Rigel Pharmaceuticals, Incorporated  
; TITLE OF INVENTION: Directed Evolution of Protein in Mammalian Cells  
; FILE REFERENCE: 021044-000110US  
; CURRENT APPLICATION NUMBER: US/10/152,296  
; CURRENT FILING DATE: 2002-12-10  
; PRIOR APPLICATION NUMBER: US 60/291,871  
; PRIOR FILING DATE: 2001-05-18  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 240  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:mammalian  
; OTHER INFORMATION: codon-optimized variant (DsRED) of Discosoma sp.  
; OTHER INFORMATION: "red" red fluorescent protein (RFP)  
US-10-152-296-2

Query Match	99.7%	Score	1210;	DB	4;	Length	240;
Best Local Similarity	99.6%	Pred. No.	5.7e-130;				
Matches	224;	Conservative	1;	Mismatches	0;	Indels	0;
						Gaps	0;
QY	1	MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKKLVTKGGPLPFAWDI	60				
Db	2	VRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKKLVTKGGPLPFAWDI	61				
QY	61	LSPOFOYGSKVYVVKHPADIPDYKKLSFPEGFKWERMNFEDGGVTVTTQDSSLQDGCIFY	120				
Db	62	LSPOFOYGSKVYVVKHPADIPDYKKLSFPEGFKWERMNFEDGGVTVTTQDSSLQDGCIFY	121				
QY	121	KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLVKGKGLKKGHLYVEFKSI	180				
Db	122	KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLVKGKGLKKGHLYVEFKSI	181				
QY	181	YMAKKPVQLPGYVYVDSKLDITSHNEDYTIQEYERTEGRHHLFL	225				
Db	182	YMAKKPVQLPGYVYVDSKLDITSHNEDYTIQEYERTEGRHHLFL	226				

RESULT 2  
US-09-277-716-16  
; Sequence 16, Application US/09277716A  
; Patent No. 6232107



```
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; APPLICANT: Szent-Gyorgyi, Christopher
; APPLICANT: PROLUME, LTD.
; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE
; CURRENT APPLICATION NUMBER: US/09/277,716A
; CURRENT FILING DATE: 1999-03-26
; EARLIER APPLICATION NUMBER: 60/102,939
; EARLIER FILING DATE: 1998-10-01
; EARLIER APPLICATION NUMBER: 60/089,367
; EARLIER FILING DATE: 1998-06-15
; EARLIER APPLICATION NUMBER: 60/079,624
; EARLIER FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Renilla mulleri
; FEATURE:
; OTHER INFORMATION: Renilla mulleri Green Fluorescent Protein (GFP)
US-09-277-716-16

Query Match      40.0%; Score 485.5; DB 3; Length 238;
Best Local Similarity 44.5%; Pred. No. 3.9e-47;
Matches 97; Conservative 43; Mismatches 71; Indels 7; Gaps 3;

QY 4 SKNVIK-----EFMRFKVRMEGTVNGHEFEIEGEGRYPYEGHNTVKLKVTKGGPLPFAW 58
   ||::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |
Db 2 SKQILKNTCLQEVMSYKVNLEGIVNNHVFTMEGCGKNILFGNQLVQIRVTKGAPLPFAF 61

QY 59 DILSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGC 118
   ||::|| ||||::: |::|| |::|| |::|| |::|| |::|| |::|| |::|| |
Db 62 DIVSPAFOYGNRTFTKYPNDISDYFIQSFPAGFMYERTLRYEDGGLVEIRSDINLIEDKF 121

QY 119 IYKVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVFE 178
   :|::: | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 122 VYRVEYKGSNFPDDGPVMQKTILGIEPSFEAMYMNNGVLVGEVILVYKLSGKYYSCHMK 181

QY 179 SIYMAKKPV-QLPGYIYVDSKLDITSHNEDYTIVEQYE 215
   : : | : | : : : : : : : : : : : : : : : : : : : : : : :
Db 182 TLMKSKGVVKEFPSYHFIOHRLEKT-YVEDGGFVEQHE 218

RESULT 3
US-09-609-161B-16
; Sequence 16, Application US/09609161B
; Patent No. 6436682
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; APPLICANT: Szent-Gyorgyi, Christopher
; APPLICANT: PROLUME, LTD.
; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE LUC
; TITLE OF INVENTION: AND FLUORESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIG
; TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS
; FILE REFERENCE: 24729-121B
; CURRENT APPLICATION NUMBER: US/09/609,161B
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/277,716
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/102,939
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 60/089,367
; PRIOR FILING DATE: 1998-06-15
; PRIOR APPLICATION NUMBER: 60/079,624
; PRIOR FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Renilla mulleri
; FEATURE:
```

```
; OTHER INFORMATION: Renilla mulleri Green Fluorescent Protein (GFP)
US-09-609-161B-16

Query Match      40.0%; Score 485.5; DB 4; Length 238;
Best Local Similarity 44.5%; Pred. No. 3.9e-47;
Matches 97; Conservative 43; Mismatches 71; Indels 7; Gaps 3;

QY 4 SKNVIK-----EFMRFKVRMEGTVNGHEFEIEGEGRYPYEGHNTVKLKVTKGGPLPFAW 58
   ||::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |
Db 2 SKQILKNTCLQEVMSYKVNLEGIVNNHVFTMEGCGKNILFGNQLVQIRVTKGAPLPFAF 61

QY 59 DILSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGC 118
   ||::|| ||||::: |::|| |::|| |::|| |::|| |::|| |::|| |::|| |
Db 62 DIVSPAFOYGNRTFTKYPNDISDYFIQSFPAGFMYERTLRYEDGGLVEIRSDINLIEDKF 121

QY 119 IYKVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVFE 178
   :|::: | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 122 VYRVEYKGSNFPDDGPVMQKTILGIEPSFEAMYMNNGVLVGEVILVYKLSGKYYSCHMK 181

QY 179 SIYMAKKPV-QLPGYIYVDSKLDITSHNEDYTIVEQYE 215
   : : | : | : : : : : : : : : : : : : : : : : : : : : : :
Db 182 TLMKSKGVVKEFPSYHFIOHRLEKT-YVEDGGFVEQHE 218

RESULT 4
US-09-626-581D-65
; Sequence 65, Application US/09626581D
; Patent No. 6548249
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; TITLE OF INVENTION: Fusions of Scaffold Proteins with Random Peptide
; FILE REFERENCE: A-66900-3/RMS
; CURRENT APPLICATION NUMBER: US/09/626,581D
; CURRENT FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 09/169,015
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 09/415,765
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 65
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Renilla muelleri
US-09-626-581D-65

Query Match      40.0%; Score 485.5; DB 4; Length 238;
Best Local Similarity 44.5%; Pred. No. 3.9e-47;
Matches 97; Conservative 43; Mismatches 71; Indels 7; Gaps 3;

QY 4 SKNVIK-----EFMRFKVRMEGTVNGHEFEIEGEGRYPYEGHNTVKLKVTKGGPLPFAW 58
   ||::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |::|| |
Db 2 SKQILKNTCLQEVMSYKVNLEGIVNNHVFTMEGCGKNILFGNQLVQIRVTKGAPLPFAF 61

QY 59 DILSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGC 118
   ||::|| ||||::: |::|| |::|| |::|| |::|| |::|| |::|| |::|| |
Db 62 DIVSPAFOYGNRTFTKYPNDISDYFIQSFPAGFMYERTLRYEDGGLVEIRSDINLIEDKF 121

QY 119 IYKVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVFE 178
   :|::: | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 122 VYRVEYKGSNFPDDGPVMQKTILGIEPSFEAMYMNNGVLVGEVILVYKLSGKYYSCHMK 181

QY 179 SIYMAKKPV-QLPGYIYVDSKLDITSHNEDYTIVEQYE 215
   : : | : | : : : : : : : : : : : : : : : : : : : : : : :
Db 182 TLMKSKGVVKEFPSYHFIOHRLEKT-YVEDGGFVEQHE 218

RESULT 5
US-09-415-765B-65
; Sequence 65, Application US/09415765B
; Patent No. 6548632
; GENERAL INFORMATION:
```









Qy	64	QFQYGSKVYVXHPADIPDYKKLSFPEGFKWERYVMNFEDGGVVTVTQDSSLQDGCFFIYKVK	123
Db	68	AFQYGNRTFTKYPDDIADYFVQSPAGFFYERNLRFEDGAIVDIRSDISLEDDKFFHYKVE	127
Qy	124	FIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLGKGEIHKALKLKDGGHYLVVEFKSIYMA	183
Db	128	YRGNGFPSNGPVMQKAILGMEPSFEVVVMNSGVLVGEVDLVYKLESGNYYSCHMKTFFRS	187
Qy	184	KKPV-QLPGYVYVDSKLDITSHNEDYTIVEQYE	215
Db	188	KGKVKEFFEYHFIHRLKKT-YVEEGSFVEQHE	219

Search completed: June 30, 2005, 15:51:10  
Job time : 44 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 1, 2005, 18:20:13 ; Search time 3340 Seconds  
(without alignments)  
2564.210 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMRFKVRMEG.....EDYTIVEQYERTEGRHHLFL 225

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US1006922/runat\_30062005\_110616\_19095/app\_query.fasta\_1.391  
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US1006922@cgn\_1\_1\_3437@runat\_30062005\_110616\_19095 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_htc:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	225	18.5	661	4	BI385436
2	216.5	17.8	637	4	BI379468
3	216.5	17.8	687	4	BI387917
4	212.5	17.5	643	4	BI386699
5	212.5	17.5	675	4	BI382638
6	209.5	17.3	790	1	AL044652
7	198.5	16.4	613	4	BI386976
8	192.5	15.9	629	4	BI388023
9	190	15.7	681	4	BI386888

10	188	15.5	629	4	BI384961	BI384961	BFLG2_003
11	187	15.4	707	4	BI382442	BI382442	BFLG2_000
12	185.5	15.3	689	4	BI380055	BI380055	BFLG1_001
13	184.5	15.2	571	4	BI387406	BI387406	BFL26_002
14	180.5	14.9	550	5	BQ551001	BQ551001	H4006G01-
15	175	14.4	671	4	BI381780	BI381780	BFLG1_003
16	173.5	14.3	561	4	BI379061	BI379061	BFLG1_000
17	172.5	14.2	577	4	BI386955	BI386955	BFL26_001
18	168.5	13.9	591	4	BI382373	BI382373	BFLG2_000
19	163	13.4	563	4	BI382999	BI382999	BFLG2_001
20	161.5	13.3	459	4	BI386952	BI386952	BFL26_001
21	158.5	13.1	537	4	BI376909	BI376909	BFLG3_000
22	131	10.8	437	4	BI388504	BI388504	BFL26_003
23	128	10.5	453	4	BI385638	BI385638	BFL26_000
24	127	10.5	629	4	BI380008	BI380008	BFLG1_001
25	122	10.0	617	4	BI381719	BI381719	BFLG1_003
26	118	9.7	541	4	BI380627	BI380627	BFLG1_002
27	111	9.1	527	4	BI381765	BI381765	BFLG1_003
28	110.5	9.1	511	1	AJ437840	AJ437840	AJ437840
29	110	9.1	397	4	BI384677	BI384677	BFLG2_002
30	107.5	8.9	675	6	CA352443	CA352443	623701_NC
31	104.5	8.6	457	4	BI378182	BI378182	BFLG3_001
32	102.5	8.4	397	1	AJ437938	AJ437938	AJ437938
33	101.5	8.4	325	4	BI386240	BI386240	BFL26_000
34	101.5	8.4	531	4	BI376416	BI376416	BFLG3_000
35	98.5	8.1	317	1	AJ437931	AJ437931	AJ437931
36	98	8.1	1302	8	CC300765	CC300765	CH261-134
37	97.5	8.0	633	4	BI376444	BI376444	BFLG3_000
38	95	7.8	697	5	BU626888	BU626888	UI-H-FT0-
39	94.5	7.8	489	4	BI381509	BI381509	BFLG1_003
40	94	7.7	553	5	BX908917	BX908917	BX908917
41	93	7.7	921	2	BF312529	BF312529	601897493
42	92.5	7.6	1573	3	CR656136	CR656136	Tetraodon
43	92	7.6	910	7	CN320807	CN320807	AGENCOURT
44	92	7.6	930	5	BQ939912	BQ939912	AGENCOURT
45	92	7.6	968	9	CG771777	CG771777	TCB45.1_F

ALIGNMENTS

RESULT 1  
BI385436  
LOCUS  
DEFINITION  
BI385436 661 bp mRNA linear EST 26-AUG-2003  
BFL26\_000139 Amphioxus 26hr cDNA library (Name convention: BFL26 or MPMGP531) Branchiostoma floridae cDNA clone MPMGP531H0446 5', mRNA sequence.  
BI385436  
BI385436.1 GI:30920419  
EST.  
Branchiostoma floridae (Florida lancelet)  
Branchiostoma floridae  
Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae; Branchiostoma.  
1 (bases 1 to 661)  
Panopoulou,G., Hennig,S., Groth,D., Krause,A., Poustka,A.J., Herwig,R., Vingron,M. and Lehrach,H.  
New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE

JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Genome Res. 13 (6A), 1056-1066 (2003)  
22683279  
12799346  
Contact: Panopoulou G  
laboratory 145, dept.Lehrach  
Max-Planck-Institut fuer Molekulare Genetik  
Inhnestr.63-73, D-14195 Berlin, Germany  
Tel: +49 30 8413 1235  
Fax: +49 30 8413 1128  
Email: panopoul@molgen.mpg.de  
The library was characterised by oligonucleotide fingerprinting (ONFP) to reduce sequencing redundancy. According to the ONFP procedure, clones giving the same hybridisation pattern with a battery of 200 8mer oligonucleotides are grouped into clusters. One



directionally cloned in pSport1 vector using a NotI  
(5'-pGACTAGTTCTAGATCGGAGCGGCCGCC (T)15-3' and a Sali 5'-  
TCGACCCACGCGTCCG-3'adapters (Gibco BRL)."

ORIGIN

Alignment Scores: 1.8e-15 Length: 637  
Pred. No.: 216.50 Matches: 53  
Score: 48.57% Conservative: 32  
Percent Similarity: 30.29% Mismatches: 85  
Best Local Similarity: 17.83% Indels: 5  
Query Match: 4 Gaps: 4  
DB:

US-10-006-922A-12 (1-225) x BI379468 (1-637)

Qy 15 LysValArgMetGluGlyThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGlu 34  
Db 79 GAGATCCACCTCCACGGTTCCATCAACGCCACGAGTTCGACCTGGTGGTGGAAA 138  
Qy 35 GlyArgProTyrGluGlyHisAsnThrValLysLeuLysValThrLysGlyGlyProLeu 54  
Db 139 GCGACCCAAACGCCGGCTCGTGGTGACACAGCGAAATCCACCAG--GGTCCCC 195  
Qy 55 PropheAlaTrpAspIleLeuSerProGlnPheGlnTyrGlySerLysValThrValLys 74  
Db 196 AAGTTCTCTCCCACTTGATGATCCCCACCTCGGGTACGGGTACTACAGTACTCCCC 255  
Qy 75 HisProAlaAspIleProAspTyrLysLysLeuSerPheProGlu-----GlyPheLys 92  
Db 256 TACCCG--GACGGACCTCGCCTTTCCAGGCCACCATGTTGGAAGATCGGGTATACA 312  
Qy 93 TrpGluArgValMetAsnPheGluAspGlyGlyValValThrValThrGlnAspSerSer 112  
Db 313 GTCCACCGCGTGTGTTGACTTCGAAGACGGAGCAAGCTGTCCATCGAGTCAAGTACTCC 372  
Qy 113 LeuGlnAspGlyCysPheIleTyrLysValLysPheIleGlyValAsnPheProSerAsp 132  
Db 373 TACGAGGTTCCCATATTAAAGCCGACATGAAGTTACAGGAAACCGGTTTCCCTGAGGAT 432  
Qy 133 GlyProValMetGlnLysLysThrMetGlyTrpGluAlaSer---ThrGluArgLeuTyr 151  
Db 433 GGGCCGGTCAAGACACGACGAGATTGTGACGACGAGCGGCTGCGTGTCCAAAGAACACCTAC 492  
Qy 152 ProArgAspGlyValLeuLysGlyGluIleHisLysAlaLeuLysLeuLysAspGlyGly 171  
Db 493 CTCAACGACACACCATCGTGGACAACTTCGACTGGACTTACAACCTGCAGATGGGAAG 552  
Qy 172 HisTyrLeuValGluPheLysSerIleTyrMetAlaLysLysPro 186  
Db 553 CGCTACAGGGCCGAGTGACGAGGCCACTATCTTCGACAAAGCCC 597

RESULT 3

BI387917  
LOCUS BI387917 687 bp mRNA linear EST 26-AUG-2003  
DEFINITION BFL26\_002620 Amphioxus 26hr cDNA library (Name convention: BFL26 or MPMGp531) Branchiostoma floridae cDNA clone MPMGp531N0767 5', mRNA sequence.

ACCESSION BI387917  
VERSION BI387917.1 GI:30922753  
KEYWORDS EST.

SOURCE Branchiostoma floridae (Florida lancelet)

ORGANISM Branchiostoma floridae  
REFERENCE 1 (bases 1 to 687)  
AUTHORS Panopoulou, G., Hennig, S., Groth, D., Krause, A., Poustka, A. J., Herwig, R., Vingron, M. and Lehrach, H.

TITLE New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes

JOURNAL Genome Res. 13 (6A), 1056-1066 (2003)  
MEDLINE 22683279  
PUBMED 12799346

COMMENT

Contact: Panopoulou G  
laboratory 145, dept. Lehrach  
Max-Planck-Institut fuer Molekulare Genetik  
ihnestr.63-73, D-14195 Berlin, Germany  
Tel: +49 30 8413 1235  
Fax: +49 30 8413 1128  
Email: panopoul@molgen.mpg.de  
The library was characterised by oligonucleotide fingerprinting (ONFP) to reduce sequencing redundancy. According to the ONFP procedure, clones giving the same hybridisation pattern with a battery of 200 8mer oligonucleotides are grouped into clusters. One clone per cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcript in the analysed library. The cluster size as well the coordinates of the rest of the clones assigned to the same fingerprint cluster as the clone from which the above EST is generated is available at the amphioxus project site at: http://www.molgen.mpg.de/amphioxus/  
Clones and filters are distributed via the Resource Center/Primary Database of the German Genome Project (http://www.rzpd.de)

PCR Primers  
FORWARD: 5' CCCCAGGCTTTACACTTATGCTTCCGGCTCG 3' (M13RSP)

BACKWARD: 5' GCTATTACGCCAGCTGGCGAAAGGGGATGTG 3' (M13FSP)

Insert Length: 1 Std Error: 0.00

Seq primer: 5'-CCGGTCCGGAATCCCGGGT-3' pSport3/86

High quality sequence stop: 687.

Location/Qualifiers

FEATURES

source  
1..687  
/organism="Branchiostoma floridae"  
/mol\_type="mRNA"  
/db\_xref="taxon:7739"  
/clone="MPMGp531N0767"  
/tissue\_type="whole embryo"  
/dev\_stage="26 hrs (neurula stage)"  
/lab\_host="E.coli, XL1 blue"  
/clone\_lib="Amphioxus 26hr cDNA library (Name convention: BFL26 or MPMGp531)"  
/note="Vector: pSport1 (Gibco BRL); Site 1: Sali, KpnI, EcoRI (5'); Site 2: NotI, BamHI, HindIII (3'); OligodT primed and directionally cloned in pSport1 vector using a NotI (5'-pGACTAGTTCTAGATCGGAGCGGCCGCC (T)15-3' and a Sali 5'- TCGACCCACGCGTCCG-3'adapters (Gibco BRL)."

ORIGIN

Alignment Scores: 2e-15 Length: 687  
Pred. No.: 216.50 Matches: 54  
Score: 48.85% Conservative: 31  
Percent Similarity: 31.03% Mismatches: 84  
Best Local Similarity: 17.83% Indels: 5  
Query Match: 4 Gaps: 4  
DB:

US-10-006-922A-12 (1-225) x BI387917 (1-687)

Qy 16 ValArgMetGluGlyThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGly 35  
Db 99 ATCCACCTCCACGGCTCCATCAACGCCACGAGTTCGACATGGTGGCGGAGGAAAGGC 158  
Qy 36 ArgProTyrGluGlyHisAsnThrValLysLeuLysValThrLysGlyGlyProLeuPro 55  
Db 159 GACCCGAAACGCCGGCTCGTGGTGACCAAGAAATCCACCAAG--GGTCCCCTGAAG 215  
Qy 56 PheAlaTrpAspIleLeuSerProGlnPheGlnTyrGlySerLysValThrValLysHis 75  
Db 216 TTCTCTCCCTACTTGATGATCCCCCCTCGGGTACGGGTACTACCAGTACCTCCCTAC 275  
Qy 76 ProAlaAspIleProAspTyrLysLysLeuSerPheProGlu-----GlyPheLysTrp 93  
Db 276 CCG---GACGACCCCTCGCCTTTCCAGGCCACCACCATGTTGGAAGGACAGGGTATGCAGTC 332  
Qy 94 GluArgValMetAsnPheGluAspGlyGlyValValThrValThrGlnAspSerSerLeu 113  
Db 333 CATCGCGTGTTCGACTTTGAAGACGGAGGCAAGCTGTCCATCGAGTTTAAGTACTCCTAC 392







US-10-006-922A-12 (1-225) x AL044652 (1-790)	
QY	16 ValArgMetGluGlyThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGly 35
Db	32 GTGGAAC TGGATGGCGATGTGAATGGGCACAAATTTCTGTCA GCGGAGAGGGTGAAGGT 91
QY	36 ArgProTyrGluGlyHisAsnThrValLysLeuLysValThrLysGlyGlyProLeuPro 55
Db	92 GATGCCACATACGGAAGCTCACCTCGAAATTCATCTGCACCACT---GGAAAGCTCCCT 148
QY	56 PheAlaTrpAspIleLeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHis 75
Db	149 GTGCCATGGCCAAACACTGGTCACTACCTTCAACCTATGGCGTGCAGTGTCTTCCAGATAC 208
QY	76 ProAlaAspIlePro-----AspTyrLysLysLeuSerPheProGluGlyPheLysTrp 93
Db	209 CCAGACCATATGAAGCAGCATGACTTTTCAAGAGCGCCATGCCCCGAGGGCTATGTGCAG 268
QY	94 GluArgValMetAsnPheGluAspGlyGlyValValThrValThrGlnAspSerSerLeu 113
Db	269 GAGAGAACCATCTTTTCAAGATGACGGGAAC TACAAGACCCCGCTGAAGTCAAGTTC 328
QY	114 GlnAspGlyCysPheIleTyrLysValLysPheIleGlyValAsnPheProSerAspGly 133
Db	329 GAAGGTGACACCCCTGGTGAATAGAATCGAGCTGAAGGCATTGACTTTAAGGAGGATGGA 388
QY	134 ProValMetGlnLysLysThrMetGlyTrpGluAlaSerThrGluArgLeuTyrProArg 153
Db	389 AACATTCTGGCCACAAAG---CTGGAATACAAC TATAACTCCCAATGTGTACATCATG 445
QY	154 AspGlyValLeuLysGlyGluIleHisLysAlaLeuLysLeuLys-----Asp 169
Db	446 GCCGACAAGCAAAAGAATGGCATCAAGGTCAACTTCAAGATCAGACACAAACATTGAGGAT 505
QY	170 GlyGlyHisTyrLeuValGluPhe-----LysSerIleTyrMetAlaLysLysProVal 187
Db	506 GGATCCGTGCAGCTGGCCGACCATTTATCAAGACAGACACTCCAATCGGCGACGGCCCTGTG 565
QY	188 GlnLeuProGlyTyrTyrValAspSerLysLeuAspIleThr-----SerHis 204
Db	566 CTCCTCCCAGACAACCATTAACCTGTCCACCCAGTCTGCCCTGTCTAAAGATCCCAACGAA 625
QY	205 AsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHisHis 222
Db	626 AAGAGAGACCACATGGTCTCTGTGGAGTTTGTGACCCGCTGTGGGATCACACAT 679
RESULT 7	
BI386976	
LOCUS	
DEFINITION	
BI386976 613 bp mRNA linear EST 26-AUG-2003	
BFL26_001679 Amphioxus 26hr cDNA library (Name convention: BFL26 or MPMGp531) Branchiostoma floridae cDNA clone MPMGp531N0799 5', mRNA sequence.	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
MEDLINE	
PUBMED	
COMMENT	
Panopoulou, G., Hennig, S., Groth, D., Krause, A., Poustka, A.J., Herwig, R., Vingron, M. and Lehrach, H. New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes Genome Res. 13 (6A), 1056-1066 (2003) 22683279 12799346 Contact: Panopoulou G Laboratory 145, dept. Lehrach Max-Planck-Institut fuer Molekulare Genetik Inhnestr.63-73, D-14195 Berlin, Germany	

Tel: +49 30 8413 1235	
Fax: +49 30 8413 1128	
Email: panopoul@molgen.mpg.de	
The library was characterised by oligonucleotide fingerprinting (ONFP) to reduce sequencing redundancy. According to the ONFP procedure, clones giving the same hybridisation pattern with a battery of 200 8mer oligonucleotides are grouped into clusters. One clone per cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcript in the analysed library. The cluster size as well the coordinates of the rest of the clones assigned to the same fingerprint cluster as the clone from which the above EST is generated is available at the amphioxus project site at: <a href="http://www.molgen.mpg.de/amphioxus/">http://www.molgen.mpg.de/amphioxus/</a>	
Clones and filters are distributed via the Resource Center/Primary Database of the German Genome Project ( <a href="http://www.rzpd.de">http://www.rzpd.de</a> )	
PCR Primers	
FORWARD: 5' CCCAGGCTTTACACTTTATGCTTCGGCTCG 3' (M13RSP)	
BACKWARD: 5' GCTATTACGCCAGCTGGCGAAAGGGGATGTG 3' (M13FSP)	
Insert length: 1 Std Error: 0.00	
Seq primer: 5'-CCGTCCTCCGAATTCCTCCGGT-3' pSport3/86	
High quality sequence stop: 613.	
FEATURES	
source	
1. .613	
/organism="Branchiostoma floridae"	
/mol_type="mRNA"	
/db_xref="taxon:7739"	
/clone="MPMGp531N0799"	
/tissue_type="whole embryo"	
/dev_stage="26 hrs (neurula stage)"	
/lab_host="E.coli, XL1 blue"	
/clone_lib="Amphioxus 26hr cDNA library (Name convention: BFL26 or MPMGp531)"	
/note="Vector: pSport1 (Gibco BRL); Site 1: Sali, KpnI, EcoRI (5'); Site 2: NotI, BamHI, HindIII (3'); OligodT primed and directionally cloned in pSport1 vector using a NotI (5'-pGACTAGTTCTAGATCGGAGCGCGCC (T)15-3' and a Sali 5'- TCGACCCACGCTCCG-3' adapters (Gibco BRL)."	
ORIGIN	
Alignment Scores:	
Pred. No.:	2.6e-13 Length: 613
Score:	198.50 Matches: 52
Percent Similarity:	47.73% Conservative: 32
Best Local Similarity:	29.55% Mismatches: 87
Query Match:	16.35% Indels: 5
DB:	4 Gaps: 4
US-10-006-922A-12 (1-225) x BI386976 (1-613)	
QY	16 ValArgMetGluGlyThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGly 35
Db	92 ATCCACCTTCACGGCTCCATCAACGCCACGAGTTCGACATGCTGGGGGAGGAAAGGC 151
QY	36 ArgProTyrGluGlyHisAsnThrValLysLeuLysValThrLysGlyGlyProLeuPro 55
Db	152 GACCCGAACGCCGGCTCGCTGGTGACCACAGCGAAATCCACCAAGGTGCC---CTGAAG 208
QY	56 PheAlaTrpAspIleLeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHis 75
Db	209 TTCTCTCCCTACTTGATGATCCCCACCTCGGGTACGGGTACTACCAGTACCTCCCTAC 268
QY	76 ProAlaAspIleProAspTyrLysLysLeuSerPheProGlu-----GlyPheLysTrp 93
Db	269 CCG---GACGGACCCCTCGCCTTTCAGGCCTCCATGTTGGAAGGATCGGGGTATGCAGTC 325
QY	94 GluArgValMetAsnPheGluAspGlyGlyValValThrValThrGlnAspSerSerLeu 113
Db	326 TACCGCGTGTTCGACTTTGAAGACGGAGGCAAGCTGACTACCGAGTTTAAGTACTCCTAC 385
QY	114 GlnAspGlyCysPheIleTyrLysValLysPheIleGlyValAsnPheProSerAspGly 133
Db	386 GAGGGTTCCCATATCAAGGCCGACATGAAGCTGATGGGAAGCGGTTTCCCTGATGACGGC 445









ORIGIN

Alignment Scores:

Pred. No.:	5.07e-12	Length:	629
Score:	188.00	Matches:	53
Percent Similarity:	43.33%	Conservative:	25
Best Local Similarity:	29.44%	Mismatches:	88
Query Match:	15.49%	Indels:	14
DB:	4	Gaps:	4

US-10-006-922A-12 (1-225) x BI384961 (1-629).

QY 3 SerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGlu----- 19

Db 60 TCATCCAAAGACCAATCGACGCAATCATGTGTCTCCGACAAACCTCGACTTGCACATC 119

QY 20 ---GlyThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyr 38

Db 120 TATGGCTCCATCAACGGTATGGAGTTCGACATGTTGGCGGTGGAGTGGCAACCCAAAG 179

QY 39 GluGlyHisAsnThrValLysLeuLysValThrLysGly-----GlyProLeu 54

Db 180 GACGGATCGCTGAGCGTAAACATGAAGTCTACCAAGGTGCCCTGCGCGTTTCTCCTCTG 239

QY 55 PropheAlaTrpAspIleLeuSerProGlnPheGlnTyrGlySerLysValTyrValLys 74

Db 240 -----CTGGTGGGCGCGCATCTGGGGTACGGCCACTATCAGTACCTACCC 284

QY 75 HisProAlaAspIleProAspTyrLysLysLeuSerPheProGluGlyPheLysTrpGlu 94

Db 285 TTCCTGACGGTCCGTCGCCCTTCCAGGACGTGTGAACAAACGGCGGTATCAATGCAT 344

QY 95 ArgValMetAsnPheGluAspGlyGlyValValThrValThrGlnAspSerSerLeuGln 114

Db 345 CGCTCTTTCAACTTCGAGACGGTGCCTGCTGACTGCCACCTACACTACTCCTACAGC 404

QY 115 AspGlyCysPheIleTyrLysValLysPheIleGlyValAsnPheProSerAspGlyPro 134

Db 405 GCGGCAAGATCCATGGAGAGTTTCATCTGTTGGGACGTGTTCCCGACGATAGTCCG 464

QY 135 ValMetGlnLysLysThrMetGlyTrpGluAlaSerThrGluArgLeuTyr---ProArg 153

Db 465 GTGATGACCAACGGCGCTGACCGTTTTGGACAGAGCGGTGGCCAAGCTGATGTGCGTGCC 524

QY 154 AspGlyValLeuLysGlyGluIleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyr 173

Db 525 GATGACAAGCTTGCCGAGTTCTGGACTGGACCTACCGCACTAGCAGCGCGGCGCTAC 584

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes

Genome Res. 13 (6A), 1056-1066 (2003)

22683279

12799346

Contact: Panopoulou G

laboratory 145, dept. Lehrach

Max-Planck-Institut fuer Molekulare Genetik

ihnestr.63-73, D-14195 Berlin, Germany

Tel: +49 30 8413 1235

Fax: +49 30 8413 1128

Email: panopoul@molgen.mpg.de

The library was characterised by oligonucleotide fingerprinting (ONFP) to reduce sequencing redundancy. According to the ONFP procedure, clones giving the same hybridisation pattern with a battery of 200 8mer oligonucleotides are grouped into clusters. One clone per cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcript in the analysed library. The cluster size as well the coordinates of the rest of the clones assigned to the same fingerprint cluster as the clone from which the above EST is generated is available at the amphioxus project site at: <http://www.molgen.mpg.de/amphioxus/>

Clones and filters are distributed via the Resource Center/Primary Database of the German Genome Project (<http://www.rzpd.de>)

PCR Primers

FORWARD: 5' CCCAGGCTTTACACTTATGCTTCGGCTCG 3' (M13RSP)

BACKWARD: 5' GCTATTACGCCAGTCGGCAAGGGGGATGTG 3' (M13FSP)

Insert Length: 1 Std Error: 0.00

Seq primer: 5'-CCGTCGCGAATTCGCGGT-3' pSport3/86

High quality sequence stop: 707.

Location/Qualifiers

1..707

/organism="Branchiostoma floridae"

/mol\_type="mRNA"

/db\_xref="taxon:7739"

/clone="MPMGp498C08131"

/tissue\_type="whole embryo"

/dev\_stage="5-6 hrs (gastrula stage)"

/lab\_host="E.coli, XL1 blue"

/clone\_lib="Amphioxus 5-6 hrs cDNA library (Name convention: BFLG or MPMGp498)"

/note="Vector: pSport1; Site\_1: Sali, KpnI, EcoRI (5'); Site\_2: NotI, BamHI, HindIII (3'); OligodT primed and directionally cloned in pSport1 vector using a NotI (5'-pGACTAGTTCTAGATCGCGAGCGCGGCC (T)15-3' and a Sali 5'-TCGACCCACGCTCCG-3'adapters (Gibco BRL)."

ORIGIN

Alignment Scores:

Pred. No.:	7.9e-12	Length:	707
Score:	187.00	Matches:	54
Percent Similarity:	42.78%	Conservative:	29
Best Local Similarity:	27.84%	Mismatches:	97
Query Match:	15.40%	Indels:	14
DB:	4	Gaps:	4

US-10-006-922A-12 (1-225) x BI382442 (1-707)

QY 3 SerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGlu----- 19

Db 89 TCATCCCAAGACCAATCGACGCAACATGTCTGTCCGACAAACCTCGACCTGCACATC 148

QY 20 ---GlyThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyr 38

Db 149 TATGGCTCCATCAACGGTATGGAGTTTGACATGTTGGCGGTGGCAACCCAAAG 208

QY 39 GluGlyHisAsnThrValLysLeuLysValThrLysGly-----GlyProLeu 54

Db 209 GACGGATCGCTGAGCGTAAACGTGAAGTCTACCAAGGTGCCCTGCGCGTTTCTCCTCTG 268

QY 55 PropheAlaTrpAspIleLeuSerProGlnPheGlnTyrGlySerLysValTyrValLys 74





sequence.	
ACCESSION	BI387406
VERSION	BI387406.1 GI:30922264
KEYWORDS	EST.
SOURCE	Branchiostoma floridae (Florida lancelet)
ORGANISM	Branchiostoma floridae
REFERENCE	Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae; Branchiostoma.
AUTHORS	1 (bases 1 to 571) Panopoulou,G., Hennig,S., Groth,D., Krause,A., Poustka,A.J., Herwig,R., Vingron,M. and Lehrach,H.
TITLE	New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes
JOURNAL	Genome Res. 13 (6A), 1056-1066 (2003)
MEDLINE	22683279
PUBMED	12799346
COMMENT	Contact: Panopoulou G Laboratory 145, dept.Lehrach Max-Planck-Institut fuer Molekulare Genetik Inhnestr.63-73, D-14195 Berlin, Germany Tel: +49 30 8413 1235 Fax: +49 30 8413 1128 Email: panopoul@molgen.mpg.de The library was characterised by oligonucleotide fingerprinting (ONFP) to reduce sequencing redundancy. According to the ONFP procedure, clones giving the same hybridisation pattern with a battery of 200 8mer oligonucleotides are grouped into clusters. One clone per cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcript in the analysed library. The cluster size as well the coordinates of the rest of the clones assigned to the same fingerprint cluster as the clone from which the above EST is generated is available at the amphioxus project site at: <a href="http://www.molgen.mpg.de/amphioxus/">http://www.molgen.mpg.de/amphioxus/</a> Clones and filters are distributed via the Resource Center/Primary Database of the German Genome Project ( <a href="http://www.rzpd.de">http://www.rzpd.de</a> ) PCR Primers FORWARD: 5' CCCAGGCTTACACTTTATGTTCCGGCTCG 3' (M13RSP) BACKWARD: 5' GCTATTAGCCAGCTGGCGAAAGGGGATGTG 3' (M13FSP) Insert Length: 1 Std Error: 0.00 Seq primer: 5'-CCGGTCCGAATTCCCGGGT-3' pSport3/86 High quality sequence stop: 571.
FEATURES	Location/Qualifiers source 1..571 /organism="Branchiostoma floridae" /mol_type="mRNA" /db_xref="taxon:7739" /clone="MPMGp531M1877" /tissue_type="whole embryo" /dev_stage="26 hrs (neurula stage)" /lab_host="E.Coli, xL1 blue" /clone_lib="Amphioxus 26hr cDNA library (Name convention: BFL26 or MPMGp531)" /note="Vector: pSport1 (Gibco BRL); Site 1: SalI, KpnI, EcoRI (5'); Site 2: NotI, BamHI, HindIII (3'); OligodT primed and directionally cloned in pSport1 vector using a NotI (5'-pGACTAGTTCTAGATCGGAGCGCGGCC (T)15-3' and a SalI 5'-TCGACCCACGCGTCCG-3'adapters (Gibco BRL)."
ORIGIN	
Alignment Scores:	
pred. No.:	1.18e-11 Length: 571
Score:	184.50 Matches: 50
Percent Similarity:	45.51% Conservative: 26
Best Local Similarity:	29.94% Mismatches: 86
Query Match:	15.20% Indels: 5
DB:	4 Gaps: 4
US-10-006-922A-12 (1-225) x BI387406 (1-571)	
QY	16 ValArgMetGluGlyThrValAsnGlyHisGluPheGluLeuGluGlyGluGlyGluGly 35
Db	77 ATCCACCTTCACGGGTCCATCAACGGCCACGAGTTCGACATGGTGGGGGAGGAAAGGC 136
QY	36 ArgProTyrGluGlyHisAsnThrValLysLeuLysValThrLysGlyGlyProLeuPro 55
Db	137 GACCCGAACGCCGCTCGTGGTGACCAACGGAATCCACCAAGGTGCC---CTGAAG 193
QY	56 PheAlaTrpAspIleLeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHis 75
Db	194 TTCTCTCCCTACTGATGATCCCCACCTCGGGTACGGGTACTACCAAGTACCTCCCTAC 253
QY	76 ProAlaAspIleProAspTyrLysLysLeuSerPheProGlu-----GlyPheLysTrp 93
Db	254 CCG---GACGGACCCCTCGCCTTTCCAGACCTCCATGTTGGAAGGATCGGGGTATGCAGTC 310
QY	94 GluArgValMetAsnPheGluAspGlyGlyValValThrValThrGlnAspSerSerLeu 113
Db	311 TACCGCGTGTTCGACTTTGAAGACGGAGGCAAGCTGTCTACCGAGTTTAAAGTACTCCTAC 370
QY	114 GlnAspGlyCysPheIleTyrLysValLysPheIleGlyValAsnPheProSerAspGly 133
Db	371 GAGGGTTCCCATATCAAGCCGACATGAAGCTGATGGGAAGCGGTTTCCCTGACGACGGC 430
QY	134 ProValMetGlnLysLysThrMetGlyTrpGluAlaSerThrGlu---ArgLeuTyrPro 152
Db	431 CCAGTCATGACCAGCCAGATTGTGGACCAGACGGGCTCGTGTCCATGAAGACGTATCTT 490
QY	153 ArgAspGlyValLeuLysGlyGluIleHisLysAlaLeuLysLeuLysAspGlyGlyHis 172
Db	491 AACGACAACACCATCGTGGACAGCTTCGACTGGAGTTACAACCTGCACNATGGGAAGCGC 550
QY	173 TyrLeuValGluPheLysSer 179
Db	551 TACAGGGCCCCGAGTGTCTCGAGC 571
RESULT 14	
BQ551001	
LOCUS	BQ551001 550 bp mRNA linear EST 20-JUN-2002
DEFINITION	H4006G01-5 NIA Mouse 7.4K cDNA Clone Set Mus musculus cDNA clone
ACCESSION	H4006G01 5', mRNA sequence.
VERSION	BQ551001
KEYWORDS	BQ551001.1 GI:21451887
SOURCE	EST.
ORGANISM	Mus musculus (house mouse)
REFERENCE	Mus musculus
AUTHORS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 550)
TITLE	VanBuren,V., Piao,Y., Dudekula,D.B., Qian,Y., Carter,M.G., Martin,P.R., Stagg,C.A., Bassey,U., Aiba,K., Hamatani,T., Kargul,G.J., Luo,A.G., Kelso,J., Hide,W. and Ko,M.S.H.
JOURNAL	Assembly, verification, and initial annotation of NIA 7.4K mouse cDNA clone set
MEDLINE	Genome Res. 12 (12), 1999-2003 (2002)
PUBMED	22354164
COMMENT	12466305 Other_ESTs: H4006G01-3 Contact: Yong Qian Laboratory of Genetics National Institute on Aging/National Institutes of Health 333 Cassell Drive, Suite 3000, Baltimore, MD 21224-6820, USA Email: <a href="mailto:cdna@lgsun.grc.nia.nih.gov">cdna@lgsun.grc.nia.nih.gov</a> This clone set has been freely distributed to the community. Please visit <a href="http://lgsun.grc.nia.nih.gov/cDNA/NIA_7_4k.html">http://lgsun.grc.nia.nih.gov/cDNA/NIA_7_4k.html</a> for details. Plate: H4006 row: G column: 01 Seq primer: -21M13 Reverse High quality sequence stop: 550 POLYA=No.
FEATURES	Location/Qualifiers source 1..550 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6" /db_xref="niaEST:H4006G01-5" /db_xref="taxon:10090"





```
QY 140 ThrMetGlyTrpGluAlaSerThrGluArgLeuTyr---ProArgAspGlyValLeuLys 158
Db 444 CTGACCACCCCTGGACAGGAGCGTGGCCAACTGCAGTGCTCGGACGACCGCACCATCACC 503
QY 159 GlyGluIleHisLysAlaLeuLysLeuLysAspGlyGly-----HisTyrLeuValGlu 176
Db 504 GGCACCAACAACCTGGAGCTTCTGCACCAACCGATGGCGATGGGAAGCGCCACCAGGCGGAT 563
QY 177 PheLysSerIleTyrMetAlaLysLysProVal 187
Db 564 GTGCAGACGAACCTACACCTTCGCCAAGCCGCTC 596
```

Search completed: July 1, 2005, 20:51:19  
Job time : 3346 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 1, 2005, 16:33:33 ; Search time 555 Seconds  
(without alignments)  
2399.895 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVIKEPMRFKVRMEG.....EDYTIVEQYTERGHHFL 225

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2 1/USPTO spool/US10006922/runat 30062005 110616 19077/app query.fasta\_1.391  
-DB=N Geneseq\_16Dec04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10006922@CGN\_1 1 470 @runat 30062005 110616 19077 -NCPV=6 -ICPV=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_16Dec04: \*  
1: geneseqn1980s: \*  
2: geneseqn1990s: \*  
3: geneseqn2000s: \*  
4: geneseqn2001as: \*  
5: geneseqn2001bs: \*  
6: geneseqn2002as: \*  
7: geneseqn2002bs: \*  
8: geneseqn2003as: \*  
9: geneseqn2003bs: \*  
10: geneseqn2003cs: \*  
11: geneseqn2003ds: \*  
12: geneseqn2004as: \*  
13: geneseqn2004bs: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1214	100.0	678	4	AAH47654
2	1214	100.0	678	4	AAD13053
3	1214	100.0	678	4	AAD11142
4	1214	100.0	678	6	ABA95905
5	1214	100.0	678	6	ABA95922

6	1214	100.0	678	6	ABA95921	Aba95921 Yeast opt
7	1214	100.0	678	6	ABA95920	Aba95920 Yeast opt
8	1214	100.0	678	6	AAD46278	Aad46278 Discosoma
9	1214	100.0	678	6	AAD28207	Aad28207 Discosoma
10	1214	100.0	678	10	ADC24127	Adc24127 Discosoma
11	1214	100.0	678	10	ADF70404	Adf70404 Discosoma
12	1214	100.0	678	11	ADL46204	Adl46204 Discosoma
13	1214	100.0	678	11	ADN33978	Adn33978 Wild-type
14	1214	100.0	678	12	ADI36420	Adi36420 Discosoma
15	1214	100.0	678	12	ADM97768	Adm97768 D sp red
16	1214	100.0	695	3	AAA48743	Aaa48743 Humanised
17	1214	100.0	859	6	AAL47952	Aal47952 Discosoma
18	1214	100.0	859	8	AAD53432	Aad53432 Discosoma
19	1214	100.0	859	10	AAD61969	Aad61969 Discosoma
20	1214	100.0	3311	9	ACA62995	Aca62995 Plasmid D
21	1210	99.7	681	10	ADC24134	Adc24134 Discosoma
22	1210	99.7	681	11	ADL46205	Adl46205 Human cod
23	1210	99.7	681	11	ADL46225	Adl46225 Human cod
24	1210	99.7	723	8	ABZ22476	Abz22476 Mammalian
25	1210	99.7	1638	10	ADL18131	Adl18131 RFP:PS(NI
26	1210	99.7	1647	10	ADL18155	Adl18155 RFP:PS(HI
27	1210	99.7	2721	13	ADS75466	Ads75466 Fibrohexa
28	1210	99.7	2772	13	ADS75468	Ads75468 Fibrohexa
29	1210	99.7	4692	6	AAL47954	Aal47954 Modified
30	1210	99.7	4692	10	ACC44640	Acc44640 Vector pD
31	1210	99.7	5436	4	AAD10003	Aad10003 Plasmid p
32	1210	99.7	6893	10	ADE24111	Ade24111 Proviral
33	1210	99.7	7910	4	AAD09979	Aad09979 pBit(dHSP
34	1210	99.7	9320	6	ABS56664	Abs56664 Plasmid p
35	1210	99.7	9658	12	ADI04104	Adi04104 Vector pT
36	1210	99.7	9678	12	ADI04103	Adi04103 Vector pT
37	1210	99.7	10263	12	ADI04102	Adi04102 Vector pT
38	1209	99.6	681	6	ABA95906	Aba95906 Yeast opt
39	1209	99.6	681	6	ABA95907	Aba95907 Yeast opt
40	1207	99.4	678	6	AAD28208	Aad28208 Discosoma
41	1205	99.3	681	4	AAH47656	Aah47656 Anthozoan
42	1203	99.1	7508	13	ADR30814	Adr30814 Zebrafish
43	1197	98.6	6985	10	ACA55359	Aca55359 Transform
44	1196	98.5	666	6	ABL61142	Ab161142 Red fluor
45	1196	98.5	711	12	ADO78067	Ado78067 Corallimo

ALIGNMENTS

RESULT 1  
AAH47654  
ID AAH47654 standard; cDNA; 678 BP.  
XX  
AC AAH47654;  
XX  
DT 30-NOV-2001 (first entry)  
XX  
DE Discosoma sp. red drFP583 protein coding sequence.  
XX  
KW Fluorescent protein; Anthozoan; fluorescence; marker; FRET; drFP583; ss.  
XX  
OS Discosoma sp.  
XX  
PN WO200162919-A1.  
XX  
PD 30-AUG-2001.  
XX  
PF 13-FEB-2001; 2001WO-US004625.  
XX  
PR 23-FEB-2000; 2000US-0184732P.  
XX  
PA (AURO-) AURORA BIOSCIENCES CORP.  
XX  
PI Nelson D, Zamaira E, Tsien R;  
XX  
DR WPI; 2001-557704/62.  
XX  
PT Proteins for Fluorescence Resonance Energy Transfer (FRET) comprise



PT functional red fluorescent proteins, and the encoding nucleic acids, with  
PT key mutations for improving the proteins function.  
XX  
PS Disclosure; Page 83; 90pp; English.

XX The invention provides a nucleic acid encoding functional red fluorescent  
CC CC protein (II) that differs from the sequence of an Anthozoan red  
CC fluorescent protein by at least one amino acid substitution, and with  
CC different fluorescent properties. The red fluorescent protein of the  
CC invention can be expressed by standard recombinant methodology. (II) are  
CC used a fluorescent markers and FRET partners. It is used for identifying  
CC protein-protein interactions. (II) is also suitable for multiplexed  
CC fluorescent analysis and FRET-based applications using existing Aequorea  
CC fluorescent proteins. (II) has improved brightness, reduced spectral  
CC cross talk, and is rapidly and efficiently expressed in mammalian cells.  
CC The key mutations in the encoding nucleic acids provide improved folding,  
CC brightness, and create (II) with sharper, more defined excitation and  
CC emission peaks when expressed in mammalian cells. The present sequence  
CC represents the coding sequence of a Discosoma sp. drFP583 protein, an  
CC anthozoan fluorescent protein

SQ Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.02e-140	Length:	678
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-10-006-922A-12 (1-225) x AAH47654 (1-678)

QY	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
Db	1	ATGAGGTCTTCCAAGATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGAAGGA	60
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
Db	61	ACGGTCAATGGGCACGAGTTTGAAATAGAGGCGGAGGAGAGGGGAGCCATACGAAGGC	120
QY	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle	60
Db	121	CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTGGGATATT	180
QY	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	181	TTGTCAACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA	240
QY	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
Db	241	GACTATAAAAAGCTGTCATTTCTCTGAAGGATTAAATGGGAAAGGGTCATGAACCTTGAA	300
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLysLeuGlnAspGlyCysPheIleTyr	120
Db	301	GACGGTGGCGTCGTTACTGTACTAACCAGGATTCCAGTTTGCAGGATGGCTGTTTCATCTAC	360
QY	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
Db	361	AAGGTCAAGTTTCATTGGCGTGAACTTTCCTTCCGATGGACCTGTTATGCAAAAGAAGACA	420
QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	421	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCTCGTGATGGCGTGTGAAAGGAGAG	480
QY	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db	481	ATTCAATAAGGCTCTGAAGCTGAAAGACGGTGGTCACTACCTAGTTGAATTCAAAAGTATT	540
QY	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	541	TACATGGCAAAAGAAGCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT	600

QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	601	ATAACAAGCCACAACGAAGACTATACAAATCGTTGAGCAGTATGAAAGAACCAGGAGCGC	660
QY	221	HisHisLeuPheLeu	225
Db	661	CACCATCTGTTCTT	675
RESULT 2			
AAD13053			
ID	AAD13053	standard; cDNA; 678 BP.	
XX			
AC	AAD13053;		
XX			
DT	16-OCT-2001	(first entry)	
XX			
DE	Discosoma sp. "red"	anthozoa fluorescent protein, drFP583 cDNA.	
XX			
KW	Protein destabilisation; linker moiety; reporter moiety; disease model;		
KW	linear multimerised domain; -NH-ubiquitin protein endoprotease;		
KW	transgenic animal; transgenic plant; disease resistance;		
KW	anthozoa fluorescent protein; natural fluorescent protein; ss.		
XX			
OS	Discosoma sp.		
XX			
PN	WO200157242-A2.		
XX			
PD	09-AUG-2001.		
XX			
PF	02-FEB-2001; 2001WO-US003791.		
XX			
PR	04-FEB-2000; 2000US-00498098.		
XX			
PA	(AURO-) AURORA BIOSCIENCES CORP.		
XX			
PI	Stack JH, Whitney M, Cubitt AB, Pollok BA;		
XX			
DR	WPI; 2001-488890/53.		
XX			
PT	Destabilizing proteins in living cells, by coupling a target protein to		
PT	linear multimerized destabilization domain non-cleavable by -NH-ubiquitin		
PT	protein endoproteases, comprising two copies of the domain.		
XX			
PS	Disclosure; Page 110; 171pp; English.		
XX			
CC	The present invention relates to a method for destabilising a target		
CC	protein in a cell. The method comprises a linker moiety which operatively		
CC	couples a target protein (a reporter moiety) to a linear multimerised		
CC	destabilising domain, which is non-cleavable by a -NH-ubiquitin protein		
CC	endoproteases. The method is useful for detecting an activity such as		
CC	protease, protein kinase or phosphoprotein phosphatase activity and is		
CC	also useful for identifying modulators of these activities. The method is		
CC	also useful for developing novel assays for a wide range of post-		
CC	translational activities, such as proteolysis, phosphorylation,		
CC	dephosphorylation, glycosylation, methylation, sulfation, prenylation,		
CC	disulfide bond formation and ADP-ribosylation within cells. The		
CC	recombinant DNA molecule of the invention is useful for creating		
CC	transgenic animals useful as disease models and transgenic plants with		
CC	improved disease resistance or other favourable traits. The present		
CC	sequence is Discosoma sp. "red" anthozoa fluorescent protein, dsFP483		
CC	cDNA which is a natural fluorescent protein used as a reporter moiety in		
CC	the exemplification of the invention		
XX			
SQ	Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;		

Alignment Scores:			
Pred. No.:	2.02e-140	Length:	678
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-10-006-922A-12 (1-225) x AAD13053 (1-678)	
QY	1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
Db	1 ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTTCGCATGGAAGGA 60
QY	21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
Db	61 ACGGTCAATGGCAGGAGTTTGAAATAGAGCGAAGGAGAGGGGAGCCATACGAAGGC 120
QY	41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60
Db	121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGGCCATTGCTTGGGATATT 180
QY	61 LeuSerProGlnPheGlnTyrGlySerLysValTyrVallLysHisProAlaAspIlePro 80
Db	181 TTGTCAACCAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCTTCCGACATACCA 240
QY	81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100
Db	241 GACTATAAAAAGCTGTCTATTTCTGAAGGATTTAAATGGGAAGGTCATGAACCTTTGAA 300
QY	101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
Db	301 GACGGTGGCGTCGTACTGTAAACCCAGGATTCAGTTTGCAGGATGGTGTTCATCTAC 360
QY	121 LysVallLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
Db	361 AAGGTCAAGTTCATTGGCGTGAACTTTCCTCCGATGGACCTGTTATGCAAAAGAAGACA 420
QY	141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
Db	421 ATGGGTGGGAAGCCAGCAGCACTGAGCGTTTGTATCCTCGTATGGCGTGTGAAAGGAGAG 480
QY	161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
Db	481 ATTCAATAAGGCTCTGAAGCTGAAAGACGGTGGTTCATTACCTAGTTGAATTCAAAAGTATT 540
QY	181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
Db	541 TACATGGCAAGAAGCCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAT 600
QY	201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
Db	601 ATAACAAGCCACAACGAAGACTATACAAATCGTTGAGCAGTATGAAGAACCAGGAGCGC 660
QY	221 HisHisLeuPheLeu 225
Db	661 CACCATCTGTTCCCT 675
RESULT 3	
AAD11142	
ID	AAD11142 standard; DNA; 678 BP.
XX	
AC	AAD11142;
XX	
DT	24-SEP-2001 (first entry)
XX	
DE	Discosoma species Anthozoa fluorescent protein, drFP583 encoding DNA.
XX	
KW	Transmembrane potential; biological membrane; fluorescent ion; detection;
KW	test chemical screening; Anthozoa fluorescent protein; FP;
KW	transgenic organism; drFP583 protein; ds.
XX	
OS	Discosoma sp.
XX	
PN	W0200142211-A2.
XX	
PD	14-JUN-2001.
XX	
PF	12-DEC-2000; 2000WO-US033739.
XX	
PR	13-DEC-1999; 99US-00459956.

XX	(REGC ) UNIV CALIFORNIA.
PA	
XX	
PI	Tsien RY, Gonzalez JE;
XX	
DR	WPI; 2001-457276/49.
XX	
PT	Determining electrical potential across a membrane in biological systems,
PT	comprises introducing two reagents, exposing the membrane to light and
PT	measuring the energy transfer.
XX	
PS	Disclosure; Page 150; 154pp; English.
XX	
CC	The patent discloses optical methods and compositions for determining
CC	transmembrane potential across biological membranes in living cells. The
CC	method of determining the electrical potential across a membrane in a
CC	biological system comprises introduction of two reagents, a first reagent
CC	comprising a hydrophobic fluorescent ion capable of redistributing from a
CC	first face of the membrane to a second face of the membrane in response
CC	to membrane potential change and a second reagent that label the first
CC	face or the second face of the membrane, which comprises a chromophore
CC	capable of undergoing energy transfer by either donating or accepting
CC	excited state energy to the fluorescent ion. The membrane is then exposed
CC	to excitation light and the energy transfer between the reagents is
CC	measured and related to the membrane potential. The method is useful for
CC	detecting changes in membrane potential in subcellular organelle
CC	membranes in biological systems. The method is used for screening of test
CC	chemicals for activity to modulate the activity of target ion channel.
CC	The invention also provides a transgenic organism comprising a first
CC	reagent that comprises a charged hydrophobic fluorescent molecule and a
CC	second reagent comprising a bioluminescent or naturally fluorescent
CC	protein. The present sequence is Discosoma species (red) DNA encoding an
CC	Anthozoa fluorescent protein (FP), drFP583
XX	
SQ	Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;
Alignment Scores:	
Pred. No.:	2.02e-140 Length: 678
Score:	1214.00 Matches: 225
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	100.00% Indels: 0
DB:	4 Gaps: 0
US-10-006-922A-12 (1-225) x AAD11142 (1-678)	
QY	1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
Db	1 ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTTCGCATGGAAGGA 60
QY	21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
Db	61 ACGGTCAATGGCAGGAGTTTGAAATAGAGCGAAGGAGAGGGGAGCCATACGAAGGC 120
QY	41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60
Db	121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGGCCATTGCTTGGGATATT 180
QY	61 LeuSerProGlnPheGlnTyrGlySerLysValTyrVallLysHisProAlaAspIlePro 80
Db	181 TTGTCAACCAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCTTCCGACATACCA 240
QY	81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100
Db	241 GACTATAAAAAGCTGTCTATTTCTGAAGGATTTAAATGGGAAGGTCATGAACCTTTGAA 300
QY	101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
Db	301 GACGGTGGCGTCGTACTGTAAACCCAGGATTCAGTTTGCAGGATGGTGTTCATCTAC 360
QY	121 LysVallLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
Db	361 AAGGTCAAGTTCATTGGCGTGAACTTTCCTCCGATGGACCTGTTATGCAAAAGAAGACA 420









especially *Escherichia coli*. RFP is useful in the same way as green fluorescent protein but is more generally applicable in modern biotechnology. (I) are optimised for expression in yeast and so generate RFP at higher levels with stronger fluorescence and thus lowers the detection limit and gives a better signal-to-noise ratio. The present sequence is that of a polynucleotide encoding the yeast optimised RFP, useful to the invention

Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

29-MAY-2002 (first entry)

Yeast optimised RFP encoding RNA SEQ ID NO 16.

Yeast; red fluorescent protein; RFP; plant; transgenic; GFP; Saccharomyces cerevisiae; Nicotiana tabacum; Arabidopsis thaliana; Escherichia coli; green fluorescent protein; biotechnology; ss.

Anthozoa.

DE20001395-U1.

15-MAR-2001.

27-JAN-2000; 2000DE-02001395.

27-JAN-2000; 2000DE-02001395.

(GPCB-) GPC BIOTECH AG.

WPI; 2002-228394/29.

New DNA encoding red fluorescent protein, useful as marker in biotechnology, has sequence optimized for expression in eukaryotes, especially yeast or plants.

Disclosure; Page 13; 19pp; German.

The invention relates to DNA (I) containing either sequence ABA95905 or sequence ABA95906 encoding a yeast optimised Red Fluorescent Protein (YRFP). (I) are used to express red fluorescent protein (RFP) in eukaryotes, especially yeast, especially Saccharomyces cerevisiae and plants, especially dicotyledonous plants including Nicotiana tabacum or Arabidopsis thaliana and also in prokaryotes, especially bacteria, especially Escherichia coli. RFP is useful in the same way as green fluorescent protein but is more generally applicable in modern biotechnology. (I) are optimised for expression in yeast and so generate RFP at higher levels with stronger fluorescence and thus lowers the detection limit and gives a better signal-to-noise ratio. The present sequence is that of an RNA sequence corresponding to the yeast optimised RFP encoding cDNA given as SEQ ID NO 1 (ABA95905)

Sequence 678 BP; 198 A; 147 C; 159 G; 0 T; 174 U; 0 Other;

Alignment Scores:

Pred. No.:	2.02e-140	Length:	678
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922A-12 (1-225) x ABA95920 (1-678)

Qy	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
Db	1	AUGAGAUUUUUAAGAACGUAUCUACAGGAAUUCUAGAGAUCAAGGUUAGAUUGGAAGGU	60
Qy	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly	40
Db	61	ACUGUUAACGGUACGAAUUCGAAUUCGAAAGGUGAGGUAAGGUAACCAUACGAAGGU	120
Qy	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle	60
Db	121	CACAACACUGUCAAAGUUAAGGUUACUAAAGGGUGGUCCAUUGCCAUCGCUUGGGACAUC	180
Qy	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	181	UUUGUCCACAAUUCCAAUACGGUUCUAAAGGUUACUAGGUGGUCCAUUGCCAUCGCUUGGGACAUC	240
Qy	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100

Db 241 GACUACAAGAGUUGUCCUCCAGAAAGGUUUAAGUGGAAAGAGUCAUGAACUUCGAA 300

QY 101 AspGlyGlyValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
|||||

Db 301 GACGGUGGUGUGUACUGUUAACUACAAGACUCCUUGCAAGACGGUUGUUAUCUAC 360  
|||||

QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
|||||

Db 361 AAGGUCAAGUUAUGGGUGUCAACUCCUUGACGGUCCAGUCAUGCAAAAGAGACU 420  
|||||

QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
|||||

Db 421 AUGGUUGGAGAGCUUCUACCGAACGUUUGUACCCCAAGACAGCGGUGUCUUGAAGGUGAA 480  
|||||

QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
|||||

Db 481 AUCCACAAGGCCUUGAAGUUGAAGGACGGUGGUCACUACUUGGUUGGAAUUCUUAUC 540  
|||||

QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrValAspSerLysLeuAsp 200  
|||||

Db 541 UACAUGGCUAAGAAGCCAGUCCAAUUGCCAGGUUACUACUACGUUGACUUAAGUUGGAC 600  
|||||

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
|||||

Db 601 AUCACCUUCUACAACGAAGACUACACUACGUUGCGAACAAUACGAACGUACUGAAGGUAGA 660  
|||||

QY 221 HisHisLeuPheLeu 225  
|||||

Db 661 CACCACUUGUUCUUG 675

RESULT 8

AAD46278

ID AAD46278 standard; DNA; 678 BP.

XX

AC AAD46278;

XX

DT 27-DEC-2002 (first entry)

XX

DE Discosoma sp. drFP583 (NFP-6) wild-type protein encoding DNA.

XX

KW Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
fluorescence activated cell sorting application; fluorescent timer;  
biosensor; fluorescence resonance energy transfer application; FRET;  
colouring agent; recombinant DNA application; analyte detection assay;  
sunscreen; second messenger detector; drFP583 protein; NFP-6; gene; ds.

XX

OS Discosoma sp.

XX

FH Key Location/Qualifiers

FT CDS 1..678

FT /\*tag= a

FT /product= "drFP583 wild-type protein"

XX

PN WO200268459-A2.

XX

PD 06-SEP-2002.

XX

PF 20-FEB-2002; 2002WO-US005749.

XX

PR 21-FEB-2001; 2001US-0270983P.

PR 04-DEC-2001; 2001US-00006922.

XX

PA (CLON-) CLONTECH LAB INC.

XX

PI Lukyanov S, Lukyanov K, Yanushevich Y, Savitsky A, Fradkov A;

XX

DR WPI; 2002-691654/74.

DR P-PSDB; AAE28833.

XX

PT New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
of an aggregating Cnidarian chromo- or fluorescent protein or mutant for  
analyte detection assays or fluorescence activated cell sorting  
applications.

PT

XX

PS Disclosure; Page 70; 80pp; English.

XX

CC The invention relates to nucleic acid molecules encoding non-aggregating  
chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
useful in analyte detection assays, as colouring agents, as markers in  
recombinant DNA applications, as sunscreens or filters, in fluorescence  
resonance energy transfer (FRET) applications, as biosensors in  
prokaryotic and eukaryotic cells, in screening assays, as second  
messenger detectors, in fluorescence activated cell sorting applications,  
in protease cleavage assays or as fluorescent timers. The present  
sequence is a DNA encoding Discosoma sp. drFP583 (NFP-6) wild-type  
protein of the invention

XX

SQ Sequence 678 BP; 146 A; 223 C; 203 G; 106 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.02e-140 Length: 678

Score: 1214.00 Matches: 225

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x AAD46278 (1-678)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
|||||

Db 1 ATGCGCTCCTCCAAGAACGTCAACAGAGATTATCGCTTCAAGGTGCGCATGGAGGC 60  
|||||

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
|||||

Db 61 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCCCTTACGAGGC 120  
|||||

QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
|||||

Db 121 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTCGCCTGGGACATC 180  
|||||

QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
|||||

Db 181 CTGTCCCCCAGTTCACAGTACGGTCCAAGGTGTACGTGAAGCACCCCGCCGACATCCC 240  
|||||

QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
|||||

Db 241 GACTACAAGAAGCTGTCTTCCCCGAGGGCTTCAAGTGGGAGCGGTGATGAACCTTCGAG 300  
|||||

QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
|||||

Db 301 GACGGCGGCTGGTGACCGTGACCCAGGACTCCTCCCTGCAGGACGGCTGCTTCATCTAC 360  
|||||

QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
|||||

Db 361 AAGGTGAAGTTCAATCGCGTGAACCTTCCCTCCGACGGCCCCGTGATGCAGAAGAGACC 420  
|||||

QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
|||||

Db 421 ATGGGCTGGGAGGCCCTCCACCGAGCGCTGTACCCCGCGACGGCGTGTGAAGGGCGAG 480  
|||||

QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
|||||

Db 481 ATCCACAAGGCCCTGAAGCTGAAGGACGGCGGCTGTACCTGGTGGAGTTCAAGTCCATC 540  
|||||

QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrValAspSerLysLeuAsp 200  
|||||

Db 541 TACATGGCCAAGAGCCCGTGCAGTGCCTCCCGCTACTACTACGTGGACTCCAAGCTGGAC 600  
|||||

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
|||||

Db 601 ATCACCTCCCAACAGGAGACTACACCATCGTGGAGCAGTACGAGCGCACCGAGGGCGGC 660  
|||||

QY 221 HisHisLeuPheLeu 225  
|||||

Db 661 CACCACCTGTTTCCTG 675

RESULT 9  
AAD28207  
ID AAD28207 standard; DNA; 678 BP.  
XX  
AC AAD28207;  
XX  
DT 22-APR-2002 (first entry)  
XX  
DE Discosoma sp. humanised wild-type Anthozoa protein drFP583 DNA.  
XX  
KW Fluorescent timer protein; protein movement; translocation; trafficking;  
KW promoter activity; gene expression; transgenic plant; gene modification;  
KW protein age; anthozoa protein; drFP583; ds.  
XX  
OS Discosoma sp.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..678  
FT /\*tag= a  
FT /product= "Humanised wild-type Anthozoa protein drFP583"  
XX  
PN WO200196373-A2.  
XX  
PD 20-DEC-2001.  
XX  
PF 13-JUN-2001; 2001WO-US019097.  
XX  
PR 14-JUN-2000; 2000US-0211607P.  
XX  
PA (CLON-) CLONTECH LAB INC.  
XX  
PI Fradkov AF, Tersikh A;  
XX  
DR WPI; 2002-154595/20.  
DR P-PSDB; AAE17540.  
XX  
PT New fluorescent timer proteins comprising an emission spectrum that  
PT changes over time from a first wavelength to a second wavelength, useful  
PT for monitoring intracellular protein movement, translocation, trafficking  
PT or stability.  
XX  
PS Example 1; Fig 1; 89pp; English.  
XX  
CC The invention relates to a fluorescent timer protein having an emission  
CC spectrum that changes over time after synthesis from a first wavelength  
CC to a second wavelength. The fluorescent timer proteins are useful in  
CC monitoring the activity of a promoter, determining the age of a protein,  
CC identifying an agent that modulates the activity of a promoter and in  
CC enriching a population of cells comprising a fluorescent timer protein.  
CC The fluorescent timer proteins are also useful for assessing gene  
CC expression during development of a multicellular organism or during  
CC cellular differentiation, in response to a drug or other inducer of  
CC promoter activity, as a reporter to serve as a read-out of promoter  
CC activity, monitoring intracellular protein movement or translocation,  
CC protein trafficking, or protein stability, to investigate temporal  
CC aspects of the activity of a regulatory element, for determining cell  
CC fate during development and organ remodelling, in spatial and temporal  
CC visualisation of newly synthesised proteins and accumulated proteases, and  
CC in distinguishing between newly formed and pre-existing structures, e.g.  
CC membrane junctions and extracellular matrix components. The fluorescent  
CC timer proteins may further be used to investigations where photobleaching  
CC techniques are employed, as detectable labels, as selectable markers, as  
CC biosensors in prokaryotic and eukaryotic cells, in protease cleavage  
CC assays, and as second messenger detectors. The nucleic acids can be used  
CC to generate transgenic, non-human plants or animals or site-specific gene  
CC modifications in cell lines. The present sequence is a DNA encoding  
CC Discosoma sp. humanised wild-type Anthozoa protein drFP583 used for  
XX generating fluorescent proteins  
SQ Sequence 678 BP; 146 A; 223 C; 203 G; 106 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.02e-140	Length:	678
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-10-006-922A-12 (1-225) x AAD28207 (1-678)			
QY	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
Db	1	ATGCGCTCCTCCAAGAACGTATCAAGGAGTTTCATCGCTTCAAGGTGCGCATGGAGGC	60
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
Db	61	ACCGTGAACGGCCACGAGTTCGAGATCGAGGCGAGGGCGAGGCGCGCCCTACGAGGC	120
QY	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle	60
Db	121	CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTCGCTTCGCCTGGACATC	180
QY	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	181	CTGTCCCCCAGTTCCAGTACGGCTCCAAGGTGTACGTGAAGCACCCCGCCGACATCCC	240
QY	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
Db	241	GACTACAAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACCTTCGAG	300
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	301	GACGGCGCGTGGTGACCGTGACCGAGGACTCCTCCCTGCAGACGGCTGCTTCATCTAC	360
QY	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
Db	361	AAGGTGAAGTTTCATCGCGTGAAGTTCCCTCCGACGGCCCCCGTGTATGCAGAAGACC	420
QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	421	ATGGGCTGGAGGCCCTCCACCGAGCGCCTGTACCCCGCGACGGCGTGTGAAGGGCGAG	480
QY	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db	481	ATCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACCTGGTGGAGTTCAAGTCCATC	540
QY	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	541	TACATGCCCAAGAAGCCCGTGCAGTCCCCGGCTACTACTACGTGGACTCCAAGTGGAC	600
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	601	ATCACCTCCCAACAGGAGACTACACCATCGTGGAGCAGTACGAGCGCACCGAGGGCCGC	660
QY	221	HisHisLeuPheLeu	225
Db	661	CACCACCTGTTCTCTG	675
RESULT 10			
ID	ADC24127	standard; DNA; 678 BP.	
XX	ADC24127;		
AC	ADC24127;		
XX	18-DEC-2003 (first entry)		
DE	Discosoma wild-type red fluorescent protein DNA #1.		
KW	Discosoma red fluorescent protein; DaRed; AB interface; AC interface;		
KW	fluorescent protein variant; transcription induction detection;		
KW	fluorescence energy resonance transfer; FRET; protein kinase;		
XX	protein phosphatase; ion indicator; ds.		
OS	Discosoma.		



XX US2003059835-A1.  
XX 27-MAR-2003.  
XX 10-APR-2002; 2002US-00121258.  
XX 26-FEB-2001; 2001US-00794308.  
PR 24-MAY-2001; 2001US-00866538.  
XX (TSIE/) TSIE R Y.  
PA (CAMP/) CAMPBELL R E.  
XX Tsien RY, Campbell RE;  
XX WPI; 2003-743764/70.  
DR P-PSDB; ADC24126.  
XX Novel polynucleotide sequence encoding Discosoma red fluorescent protein  
PT variant having a reduced propensity to oligomerize, useful for detecting  
PT transcriptional activity.  
XX Example 1; SEQ ID NO 2; 67pp; English.  
CC The invention describes a polynucleotide sequence (I) encoding a  
CC Discosoma red fluorescent protein (DsRed) variant having a reduced  
CC propensity to oligomerize, comprising amino acid substitutions at the AB  
CC and/or AC interfaces of the wild-type DsRed sequence (S1) comprising 225  
CC amino acids, given in the specification, where the substitutions result  
CC in reduced propensity of the DsRed variant to form tetramers. (I) is  
CC useful for detecting transcriptional activity by providing a host cells  
CC conatinig a vector which comprises (I) operatively linked to an  
CC expression control sequence, and an unit to assay the variant fluorescent  
CC protein fluorescence, and assaying fluorescence of the variant  
CC fluorescent protein produced by (VII), where variant fluorescent protein  
CC fluorescence is indicative of transcriptional activity. A polynucleotide  
CC encoding a fusion protein is useful for the analysis of in vivo  
CC localisation or trafficking of a polypeptide of interest. A polypeptide  
CC marker is useful as markers to identify the location and amount of a  
CC target protein produced, where the target protein is fused to the marker,  
CC as a complement to or alternative for the green fluorescent protein or  
CC its spectral variant, for detecting induction of transcriptions, in  
CC applications involving fluorescence energy resonance transfer (FRET),  
CC which detects events as the function of the movement of fluorescent  
CC donors and acceptors towards or away from each other, for making  
CC fluorescent sensors for protein kinase and phosphatase activities or  
CC indicators for ions and molecules such as Ca2+, Zn2+, for identifying the  
CC presence of a molecule in a sample, for identifying a specific  
CC interaction of a first and second molecule, for determining whether a  
CC sample contains an enzyme or for determining the pH of the sample. (I) is  
CC useful for identifying a region or condition that regulates the activity  
CC of an expression control sequence. This sequence encodes Discosoma wild-  
XX type red fluorescent protein.  
SQ Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.02e-140 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-10-006-922A-12 (1-225) x ADC24127 (1-678)

Qy 1 MetArgSerSerLysAsnValileLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 1 ATGAGGCTCTCCAGAAGATGTTATCAAGGAGTTCATGAGGTTTAAAGGTTGCGATGGAAGGA 60  
Qy 21 ThrValAsnGlyHisGluPheGluileGluGlyGluGlyGluGlyArgProTyrGluGly 40  
Db 61 ACGGTCAATGGGACGAGGTTTGAATAGAGGCGAAGGAGAGGGGGAGCCATACGAAGGC 120

QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 121 CACAATACCGTAAAGCTTAAGGTAAACCAAGGGGGACCTTTGCCATTGCTTGGGATATT 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 181 TTGTCAACCAAAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCTGCCGACATACCA 240  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
Db 241 GACTATAAAAAAGCTGTCTATTTCTGAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAA 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 361 AAGGTCAAGTTCATTTGGCGTGAACCTTCTCCGATGGACCTGTTATGCAAAAGAAGACA 420  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 481 ATTCAATAAGGCTCTGAAGCTGAAGACGGTGGTGCATTACCTAGTTGAATTCAAAAAGTATT 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 541 TACATGGCAAGAAGCCCTGTGCAGCTACCAGGTACTACTATGTTGACTCCAAACTGGAT 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCAGGAGGACGC 660  
QY 221 HisHisLeuPheLeu 225  
Db 661 CACCATCTGTTTCCTT 675  
RESULT 11  
ADF70404  
ID ADF70404 standard; cDNA; 678 BP.  
XX ADF70404;  
AC ADF70404;  
XX 12-FEB-2004 (first entry)  
XX Discosoma wild-type GFP variant cDNA SeqID27.  
DE ligand; orphan receptor protein; fusion protein; fluorescent protein;  
KW cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;  
KW GFPuv; Enhanced GFP; EGFP; gene; ss.  
XX Discosoma sp.  
OS WO2003071272-A1.  
XX 28-AUG-2003.  
PD 21-FEB-2003; 2003WO-JP001901.  
XX 22-FEB-2002; 2002JP-00045728.  
PR 23-JUL-2002; 2002JP-00213949.  
PR 11-OCT-2002; 2002JP-00298237.  
XX (TAKE ) TAKEDA CHEM IND LTD.  
PA Hinuma S, Fujii R, Ogi K, Komatsu H, Kawamata Y, Hosoya M;  
XX WPI; 2003-697654/66.  
DR P-PSDB; ADF70403.



XX Transformation of cells with a fusion protein of an orphan receptor  
PT protein with a fluorescent protein useful for identification of ligands  
PT to the orphan receptor.  
XX  
PS Disclosure; SEQ ID NO 27; 594pp; Japanese.  
XX  
CC This invention relates to a novel method of identifying ligands to an  
CC orphan receptor protein which comprises transforming cells with DNA  
CC encoding a fusion protein of the orphan receptor with a fluorescent  
CC protein, so that the fusion protein is expressed in the cells (or cell  
CC membranes isolated from them) and contacting the cells with the potential  
CC ligand to be tested. A suitable fluorescent protein for incorporation in  
CC the fusion protein is green fluorescent protein (GFP), for example GFP-1,  
CC wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the  
CC identification of ligands binding to an orphan receptor protein.  
XX  
SQ Sequence 678 BP; 204 A; 129 C; 179 G; 166 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.02e-140 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-10-006-922A-12 (1-225) x ADF70404 (1-678)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 1 ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTCATGAGTTTAAGGTTGCGATGGAAGGA 60

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 61 ACGGTCAATGGGCACGAGTTTGAAATAGAGCGAAGGAGAGGGGAGGCCATACGAAGGC 120

QY 41 HisAsnThrValLysLeuLysValThrLysGlyProLeuProPheAlaTrpAspIle 60  
Db 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTCCTTGGGATATT 180

QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrVallysHisProAlaAspIlePro 80  
Db 181 TTGTCACCAACAATTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240

QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
Db 241 GACTATAAAAAGTGTCATTTCCTGAAGGATTTAAATGGGAAAGGTCATGAACTTTGAA 300

QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGGTGGCGTCTACTGTACCCAGGATTCAGATTTCAGGATGGCTGTTTCATCTAC 360

QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 361 AAGGTCAAGTTTCATTGGCGTGAACTTTCCCTCCGATGGACCTGTTATGCAAAAGAAGACA 420

QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 421 ATGGGCTGGGAAGCCAGCACCTGAGCGTTTGATCTCCTCGTATGGCGTGTGAAAGGAGAG 480

QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATCAAAAGTATT 540

QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 541 TACATGGCAAGAAGCCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAT 600

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATAACAAGCCACACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGGAGCGC 660

QY 221 HisHisLeuPheLeu 225  
Db 661 CACCATCTGTTCCTT 675

RESULT 12  
ADL46204  
ID ADL46204 standard; DNA; 678 BP.  
XX  
AC ADL46204;  
XX  
DT 20-MAY-2004 (first entry)  
XX  
DE Discosoma red fluorescent protein (DsRed) coding sequence.  
XX  
KW ds; gene; red fluorescent protein; DsRed; fluorescence; red wavelength;  
KW oligomerization; tetramerization; immunoassay; hybridization assay.  
XX  
OS Discosoma sp.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..678  
FT /\*tag= a  
FT /product= "DsRed protein"  
XX  
PN WO2003086446-A1.  
XX  
PD 23-OCT-2003.  
XX  
PF 09-APR-2003; 2003WO-US010879.  
XX  
PR 10-APR-2002; 2002US-00121258.  
PR 29-JUL-2002; 2002US-00209208.  
XX  
PA (REGC ) UNIV CALIFORNIA.  
XX  
PI Tsien RY, Campbell RE, Baird GS;  
XX  
DR WPI; 2003-845265/78.  
DR P-PSDB; ADL46203.  
XX  
PT New monomeric and dimeric Anthozoan fluorescent protein variants with  
PT reduced propensity to oligomerize, and encoding polynucleotides, useful  
PT in molecular biology, e.g. in immunoassays or in tracking protein  
PT movement in cells.  
XX  
PS Disclosure; SEQ ID NO 2; 166pp; English.  
XX  
CC The invention relates to a polynucleotide sequence encoding a Discosoma  
CC red fluorescent protein (DsRed) variant having a reduced propensity to  
CC oligomerize. The protein variant comprises one or more amino acid  
CC substitutions at the AB and/or AC interface(s) of the wild-type DsRed  
CC sequence, where the substitutions result in reduced propensity of the  
CC DsRed variant to form tetramers and where the variant displays detectable  
CC fluorescence of at least one red wavelength. The composition and methods  
CC are useful in producing red fluorescent proteins having reduced  
CC propensity for oligomerization, especially tetramerization. The protein  
CC may be used in molecular biology and in other scientific applications,  
CC such as in immunoassays or hybridization assays, or in tracking the  
CC movement of proteins in cells. This sequence corresponds to the DsRed  
CC coding sequence.  
XX  
SQ Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.02e-140 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 11 Gaps: 0

US-10-006-922A-12 (1-225) x ADL46204 (1-678)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
DB 1 ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTGCGATGGAAGGA 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 61 ACGGTCAATGGGACGAGTTTGAATAGAAAGCGAAGGAGAGGGAGGCCATACGAAGGC 120  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 121 CACAATACCGTAAAGCTTAAGGTAAACCAAGGGGACCTTTGCCATTGCTTGGGATATT 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
DB 181 TTGTCACCAACAATTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
DB 241 GACTATAAAAAGCTGTCATTTCTGAAGGATTTAAATGGAAAGGGTCATGAACCTTTGAA 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGTGGCGTCTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
DB 361 AAGGTCAAGTTTCATTTGGCGTGAACCTTTCCTCCGATGGACCTGTTATGCAAAAGACACA 420  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
DB 421 ATGGGCTGGGAAGCCAGCAGCTAGCGGTTTGTATCCTCGTGTATGGCGTGTGAAAGGAGAG 480  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyHisTyrLeuValGluPheLysSerIle 180  
DB 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCACTTACCTAGTTGAATCAAAAGTATT 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
DB 541 TACATGGCAAGAAAGCCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAT 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAAGCCGAGGGACGC 660  
QY 221 HisHisLeuPheLeu 225  
DB 661 CACCATCTGTTCTT 675

RESULT 13

ADN33978

ID ADN33978 standard; DNA; 678 BP.

XX AC ADN33978;

XX DT 01-JUL-2004 (first entry)

XX DE Wild-type DsRED encoding sequence.

XX KW Cnidarian; fluorescence resonance energy transfer; FRET; wild-type DsRED;

XX OS DsRED.

XX PN WO2003054158-A2.

XX PD 03-JUL-2003.

XX PF 18-DEC-2002; 2002WO-US040539.

XX PR 19-DEC-2001; 2001US-0341723P.

XX PA (UYCH-) UNIV CHICAGO.

PI Bevis B, Glick B;  
XX WPI; 2003-569236/53.  
DR P-PSDB; ADN33979.  
XX Novel nucleic acid encoding a rapidly maturing chromo- or fluorescent mutant of a Cnidarian chromo- or fluorescent protein or its mutant, useful for applications involving chromo- or fluorescent proteins.  
PS Claim 8; SEQ ID NO 1; 65pp; English.  
XX The present invention relates to nucleic acid that encodes a rapidly maturing chromo or fluorescent mutant of a Cnidarian chromo- or fluorescent protein or its mutant. The protein is useful in applications involving nucleic acid encoding a chromo- or fluorescent protein and is useful for producing a chromo and/or fluorescent protein which involves growing the cell, whereby the protein is expressed, and isolating the protein substantially free of other proteins. The protein is useful in applications involving chromo- or fluorescent protein and is useful as PCR primers, hybridization probes, etc. The expression cassettes are useful for synthesizing related proteins. The chromoproteins are useful as coloring agents which are capable of imparting color or pigment to a particulate composition of matter e.g. food compositions, pharmaceuticals, cosmetics, living organisms, e.g., animals and plants. The chromoproteins may also find use as labels in analyte detection assays, e.g. assays for biological analytes of interest and as selectable markers in recombinant DNA applications, e.g. the production of transgenic cells and organisms. The fluorescent proteins find use in a variety of different applications, e.g. in fluorescence resonance energy transfer (FRET) applications, as biosensors in prokaryotic and eukaryotic cells, in applications involving the automated screening of arrays of cells expressing fluorescent reporting groups by using microscopic imaging and electronic analysis, as second messenger detectors, and in fluorescence activated cell sorting applications and as in vivo marker in animals. The fluorescent proteins also find use in protease cleavage assays. The proteins can also be used is assays to determine the phospholipid composition in biological membranes and as a fluorescent timer. The present sequence represents the wild-type DsRED encoding sequence.  
SQ Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.02e-140 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 11 Gaps: 0

US-10-006-922A-12 (1-225) x ADN33978 (1-678)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20

DB 1 ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTGCGATGGAAGGA 60

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40

DB 61 ACGGTCAATGGGACGAGTTTGAATAGAAAGCGAAGGAGAGGGGAGGCCATACGAAGGC 120

QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60

DB 121 CACAATACCGTAAAGCTTAAGGTAAACCAAGGGGACCTTTGCCATTGCTTGGGATATT 180

QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80

DB 181 TTGTCACCAACAATTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240

QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100

DB 241 GACTATAAAAAGCTGTCATTTCTGAAGGATTTAAATGGAAAGGGTCATGAACCTTTGAA 300

QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120

Db 301 GACGGTGGCGTCGTTACTGTAAACCCAGGATTCCAGTTTGCAGGATGGCTGTTTCATCTAC 360

Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
|||||

Db 361 AAGGTCAAGTTCATTGGCGTGAACTTTCCTCCGATGGACCTGTTATGCAAAAGAGACA 420

Qy 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
|||||

Db 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAGGAGAG 480

Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
|||||

Db 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT 540

Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
|||||

Db 541 TACATGGCAAGAAGCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600

Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
|||||

Db 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCGAGGGACGC 660

Qy 221 HisHisLeuPheLeu 225

Db 661 CACCATCTGTTCCTT 675

RESULT 14  
ADI36420

ID ADI36420 standard; DNA; 678 BP.

XX

AC ADI36420;

XX

DT 22-APR-2004 (first entry)

XX

DE Discosoma sp. red fluorescent protein (RED) DNA.

XX

KW Fluorobody; binding ligand; green fluorescent protein; GFP;

KW target detection; gene; red fluorescent protein; ds; RED.

XX

OS Discosoma sp.

XX

FH Key Location/Qualifiers

FT CDS 1..678

FT /\*tag= a

FT /product= "Discosoma sp. red fluorescent protein (RED)"

XX

PN US2003203355-A1.

XX

PD 30-OCT-2003.

XX

PF 24-APR-2002; 2002US-00132067.

XX

PR 24-APR-2002; 2002US-00132067.

XX

PA (LALA-) LOS ALAMOS NAT LAB.

PA (REGC ) UNIV CALIFORNIA.

XX

PI Bradbury AM, Zeytun A, Waldo GS;

XX

DR WPI; 2004-154325/15.

DR P-PSDB; ADI36421.

XX

PT Novel binding ligand with intrinsic fluorescence and comprising

PT fluorescent protein having heterologous binding sites, useful for

PT detecting target molecule.

XX

PS Example 6; SEQ ID NO 3; 23pp; English.

XX

CC The invention relates to binding ligands (fluorobodies) with intrinsic

CC fluorescence, which comprises green fluorescent protein (GFP) having

CC heterologous binding sites. The binding ligand is useful for detecting

CC the target molecule and is efficiently detects the target molecule. The

CC present sequence is Discosoma sp. red fluorescent protein (RED) DNA used

CC in the exemplification of the invention.

XX

SQ Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.02e-140 Length: 678

Score: 1214.00 Matches: 225

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 12 Gaps: 0

US-10-006-922A-12 (1-225) x ADI36420 (1-678)

Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
|||||

Db 1 ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTTCCATGGAAGGA 60

Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
|||||

Db 61 ACGGTCAATGGGCACGAGTTTGAATAGAAAGCGAAGGAGAGGGGAGGCCATACGAAGGC 120

Qy 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
|||||

Db 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTGCTTGGGATATT 180

Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
|||||

Db 181 TTGTCAACCAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTCCGACATACCA 240

Qy 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
|||||

Db 241 GACTATAAAAAGCTGTCATTTCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300

Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
|||||

Db 301 GACGGTGGCGTCGTTACTGTAAACCAGGATTCAGTTTGCAGATGGCTGTTTCATCTAC 360

Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
|||||

Db 361 AAGGTCAAGTTCATTGGCGTGAACCTTTCCTCCGATGGACCTGTATATGCAAAAGAGACA 420

Qy 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
|||||

Db 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAGGAGAG 480

Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
|||||

Db 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT 540

Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
|||||

Db 541 TACATGGCAAGAAGCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600

Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
|||||

Db 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCGAGGGACGC 660

Qy 221 HisHisLeuPheLeu 225

Db 661 CACCATCTGTTCCTT 675

RESULT 15  
ADM97768

ID ADM97768 standard; DNA; 678 BP.

XX

AC ADM97768;

XX

DT 01-JUL-2004 (first entry)

XX

DE D sp red fluorescent protein coding sequence SEQ ID NO: 21.

XX

KW ds; gene; enzyme; sensor cell; fluorescent protein;



KW signal transduction detection system; promoter; targeting sequence;  
KW targeted drug.

XX Discosoma sp.

FH Key Location/Qualifiers  
FT CDS 1..678  
FT /\*tag= a  
FT /product= "fluorescent protein"  
XX

PN WO2004031415-A2.

XX

PD 15-APR-2004.

XX

PF 05-SEP-2003; 2003WO-US028078.

XX

PR 05-SEP-2002; 2002US-0408297P.

XX

PA (VERT-) VERTEX PHARM INC.

XX

PI Whitney MA, Zeh K, Sanders PS;

XX

DR WPI; 2004-330208/30.

DR P-PSDB; ADM97769.

XX

PT Developing a sensor cell, useful in determining the activity of a target  
PT gene and in developing therapeutic drugs, comprises providing cells  
PT comprising a signal transduction detection system and introducing DNA  
PT construct into cells.

XX

PS Disclosure; Page 167-168; 234pp; English.

XX

CC The present invention relates to a method of developing a sensor cell,  
CC for determining the activity of a target gene in the cell, which  
CC comprises providing a homogeneous population of cells, where each of the  
CC cells comprises a signal transduction detection system and introducing  
CC into the population of cells an isolated DNA construct comprising a  
CC promoter operatively linked to a targeting sequence. The method is useful  
CC in developing a sensor cell for determining the activity of a target gene  
CC in the cell. The sensor cell and the methods are useful in developing new  
CC and therapeutic drugs directed to the targets. The present sequence is a  
CC coding sequence shown in the exemplification of the invention.

XX

SQ Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.02e-140 Length: 678

Score: 1214.00 Matches: 225

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 12 Gaps: 0

US-10-006-922A-12 (1-225) x ADM97768 (1-678)

Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20

Db 1 ATGAGGTCCTCCAAGAAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTTCGCATGGAAGGA 60

Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40

Db 61 ACGGTCAATGGGCACGAGTTTGAATAGAGCGGAGGAGAGGGAGGCCATACGAAGGC 120

Qy 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60

Db 121 CACAATACCGTAAGCTTAAGGTAAAGGTAACCAAGGGGGACCTTTGCCATTGCTTGGGATATT 180

Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80

Db 181 TTGTACCAACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCTGCCGACATACCA 240

Qy 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100

Db 241 GACTATAAAAAAGCTGTCAATTTCTCTGAAGGATTTAAATGGGAAAGGCTCATGAACTTTGAA 300

Qy 101 AspGlyGlyValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120

Db 301 GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360

Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140

Db 361 AAGGTCAAGTTCATTGGCGTGAACCTTCCTCCGATGGACCTGTTATGCAAAAGAAGACA 420

Qy 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160

Db 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGATATCCTCGTGATGGCGTGTGAAAGGAGAG 480

Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180

Db 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTGCATTACCTAGTTGAATTCAAAAAGTATT 540

Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrValAspSerLysLeuAsp 200

Db 541 TACATGGCAAAAGAAGCCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600

Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220

Db 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCAGGAGCGC 660

Qy 221 HisHisLeuPheLeu 225

Db 661 CACCATCTGTTCTCTT 675

Search completed: July 1, 2005, 18:42:03  
Job time : 558 secs



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 1, 2005, 18:32:43 ; Search time 614 Seconds  
(without alignments)  
2298.439 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMRFKVRMEG.....EDYTIVEQYERTEGRHHLFL 225

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 6313374 seqs, 3136092125 residues

Total number of hits satisfying chosen parameters: 12626748

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US1006922/runat\_30062005\_110617\_19138/app\_query.fasta\_1.391  
-DB=Published\_Applications\_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosu62  
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US1006922@cgn\_1\_1\_480@runat\_30062005\_110617\_19138  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq.\*
- 20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq.\*
- 21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq.\*
- 22: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq.\*
- 24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1214	100.0	678	9	US-09-967-772-6	Sequence 6, Appli
2	1214	100.0	678	13	US-10-006-922-11	Sequence 11, Appl
3	1214	100.0	678	14	US-10-081-864-7	Sequence 7, Appli
4	1214	100.0	678	14	US-10-121-258-2	Sequence 2, Appli
5	1214	100.0	678	16	US-10-315-920-1	Sequence 1, Appli
6	1214	100.0	678	17	US-10-132-067-3	Sequence 3, Appli
7	1214	100.0	678	17	US-10-335-517-6	Sequence 6, Appli
8	1214	100.0	678	17	US-10-334-288-6	Sequence 6, Appli
9	1214	100.0	678	19	US-10-311-030-5	Sequence 5, Appli
10	1214	100.0	678	21	US-10-656-029-21	Sequence 21, Appl
11	1214	100.0	678	21	US-10-505-486-27	Sequence 27, Appl
12	1214	100.0	859	9	US-09-999-745-66	Sequence 66, Appl
13	1214	100.0	859	10	US-09-866-538-11	Sequence 11, Appl
14	1214	100.0	859	10	US-09-794-308-11	Sequence 11, Appl
15	1214	100.0	859	10	US-09-865-291-11	Sequence 11, Appl
16	1214	100.0	859	19	US-10-433-640-12	Sequence 12, Appl
17	1214	100.0	859	20	US-10-885-988-11	Sequence 11, Appl
18	1214	100.0	859	21	US-10-857-622-11	Sequence 11, Appl
19	1214	100.0	3311	10	US-09-797-496B-3	Sequence 3, Appli
20	1210	99.7	681	13	US-10-006-922-35	Sequence 35, Appl
21	1210	99.7	681	14	US-10-121-258-3	Sequence 3, Appli
22	1210	99.7	681	14	US-10-121-258-23	Sequence 23, Appl
23	1210	99.7	681	19	US-10-311-030-8	Sequence 8, Appli
24	1210	99.7	713	19	US-10-311-030-11	Sequence 11, Appl
25	1210	99.7	713	19	US-10-311-030-12	Sequence 12, Appl
26	1210	99.7	723	14	US-10-152-296-1	Sequence 1, Appli
27	1210	99.7	723	19	US-10-739-656-1	Sequence 1, Appli
28	1210	99.7	1638	15	US-10-214-932-51	Sequence 51, Appl
29	1210	99.7	1647	15	US-10-214-932-75	Sequence 75, Appl
30	1210	99.7	4692	15	US-10-161-403-29	Sequence 29, Appl
31	1210	99.7	4692	19	US-10-433-640-16	Sequence 16, Appl
32	1210	99.7	5436	21	US-10-169-050-46	Sequence 46, Appl
33	1210	99.7	6984	13	US-10-001-189-45	Sequence 45, Appl
34	1210	99.7	7910	21	US-10-169-050-20	Sequence 20, Appl
35	1210	99.7	9320	19	US-10-471-065-20	Sequence 20, Appl
36	1210	99.7	9658	19	US-10-609-019-4	Sequence 4, Appli
37	1210	99.7	9678	19	US-10-609-019-3	Sequence 3, Appli
38	1210	99.7	10263	19	US-10-609-019-2	Sequence 2, Appli
39	1207	99.4	678	16	US-10-315-920-3	Sequence 3, Appli
40	1203	99.1	7508	19	US-10-742-828-4	Sequence 4, Appli
41	1198	98.7	681	13	US-10-006-922-37	Sequence 37, Appl
42	1196	98.5	666	19	US-10-332-733-22	Sequence 22, Appl
43	1196	98.5	711	19	US-10-314-936-1	Sequence 1, Appli
44	1196	98.5	711	19	US-10-314-936-3	Sequence 3, Appli
45	1196	98.5	711	23	US-11-021-014-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-967-772-6  
; Sequence 6, Application US/09967772  
; Patent No. US20020164577A1  
; GENERAL INFORMATION:  
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: TSJEN, Roger  
; APPLICANT: GONZALEZ, Jesus  
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY OPTICAL METHODS  
; FILE REFERENCE: REGEN1290-5  
; CURRENT APPLICATION NUMBER: US/09/967,772  
; CURRENT FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: US 09/459,956  
; PRIOR FILING DATE: 1999-12-13  
; PRIOR APPLICATION NUMBER: US 08/765,860  
; PRIOR FILING DATE: 1996-12-19

; PRIOR APPLICATION NUMBER: PCT/ US96/09652
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: US 08/481,977
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Discosoma sp "red"
US-09-967-772-6

Alignment Scores:
Pred. No.: 4.4e-143 Length: 678
Score: 1214.00 Matches: 225
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-006-922A-12 (1-225) x US-09-967-772-6 (1-678)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
Db 1 ATGAGGCTCTCCAAGAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTGCGATGGAAGGA 60
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
Db 61 ACGGTCAATGGGCACGAGTTTGAAATAGAGCGGAAGGAGAGGGGAGCCATACGAAGGC 120
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60
Db 121 CACAATACCGTAAGCTTAAGTTAACCAAGGGGGACCTTTGCCATTGCTTGGGATATT 180
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80
Db 181 TTGTCAACCACAAATTCAGTATGGAAGCAAGGTATATGTCAAGCACCTTGCCGACATACCA 240
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100
Db 241 GACTATAAAAAGCTGTGCATTTCTCTGAAGGATTTAAATGGGAAGGGTTCATGAACCTTTGAA 300
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
Db 301 GACGGTGGCGTCGTACTGTAAACCCAGGATTCAGTTTGACAGGATGGCTGTTTCATCTAC 360
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
Db 361 AAGGTCAAGTTCATTTGGCGTGAACTTTTCCTTCGATGGACCTGTTATGCCAAAAGAGACA 420
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
Db 421 ATGGGCTGGGAAGCCAGCACACTGAGCGTTTGATATCCTCGTGATGGCGTGTGAAAGGAGAG 480
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyHisTyrLeuValGluPheLysSerIle 180
Db 481 ATTCATAAGGCTCTGAAGCTGAAGACGGTGGTTCATTACCTAGTTGAATCAAAAAGTATT 540
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
Db 541 TACATGGCAAAAGAGCCCTGTGCAGCTACCCAGGTACTACTATGTTGACTCCAAACTGGAT 600
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
Db 601 ATAACAAGCCACACGAAGACTATACAATCGTTGAGCAGTATGAAAGACCGGAGCGC 660
QY 221 HisHisLeuPheLeu 225
Db 661 CACCATCTGTTTCCTT 675

RESULT 2
US-10-006-922-11
; Sequence 11, Application US/10006922

; Publication No. US20020197676A1
; GENERAL INFORMATION:
; APPLICANT: Lukyanov, Sergey A
; APPLICANT: Fradkov, Arcady F.
; APPLICANT: Labas, Yulii A.
; APPLICANT: Matz, Mikhail V.
; APPLICANT: Tersikh, Alexey
; TITLE OF INVENTION: No. US20020197676A1el Chromophores/Fluorophores and
; FILE REFERENCE: CLON-035CIP
; CURRENT APPLICATION NUMBER: US/10/006,922
; CURRENT FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 09/120,330
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/457,898
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,144
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,477
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/457,556
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/444,338
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Discosoma species
US-10-006-922-11

Alignment Scores:
Pred. No.: 4.4e-143 Length: 678
Score: 1214.00 Matches: 225
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-006-922-11 (1-678)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
Db 1 ATGAGGCTCTCCAAGAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTGCGATGGAAGGA 60
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
Db 61 ACGGTCAATGGGCACGAGTTTGAAATAGAAGCGGAAGGAGAGGGGAGCCATACGAAGGC 120
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60
Db 121 CACAATACCGTAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTGCTTGGGATATT 180
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80
Db 181 TTGTCAACCACAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCTTGCCGACATACCA 240
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100
Db 241 GACTATAAAAAGCTGTGCATTTCTCTGAAGGATTTAAATGGGAAGGGTTCATGAACCTTTGAA 300
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
Db 301 GACGGTGGCGTCGTACTGTAAACCCAGGATTCAGTTTGACAGGATGGCTGTTTCATCTAC 360
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
Db 361 AAGGTCAAGTTCATTTGGCGTGAACCTTCCTTCGATGGACCTGTATGCAAAAGAGACA 420
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
Db 421 ATGGGCTGGGAAGCCAGCACACTGAGCGTTTGATATCCTCGTGATGGCGTGTGAAAGGAGAG 480





Db 61 ACGGTCAATGGGCACGAGTTTGAAATAGAAAGCGGAAGGAGAGAGGGAGGCCATACGAAGGC 120  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 181 TTGTCAACCAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA 240  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
Db 241 GACTATAAAAGCTGTCATTTCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGGTGGCGTCTGTTACTGTAAACCCAGGATTCAGATTGCAGGATGGCTGTTTCATCTAC 360  
QY 121 LysValLysPheIleGlyValAlaAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 361 AAGGTCAAGTTTCAATGGCGTGAACCTTTCCCTCCGATGGACCTGTTATGCAAAAGAGACA 420  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 421 ATGGCTGGGAAGCCAGCACTGAGCGTTTGATCTCGTGATGGCGTGTGAAAGGAGAG 480  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATCAAAAGTATT 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 541 TACATGGCAAAGAGCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATAACAAGCCACACAAGAGACTATACAATCGTTGAGCAGTATGAAAGAACCGGAGCGC 660  
QY 221 HisHisLeuPheLeu 225  
Db 661 CACCATCTGTTTCCTT 675

RESULT 5

US-10-315-920-1  
; Sequence 1, Application US/10315920  
; Publication No. US20030175809A1  
; GENERAL INFORMATION:  
; APPLICANT: Fradkov, Arcady Fedorovich  
; APPLICANT: Terskikh, Alexey  
; TITLE OF INVENTION: FLUORESCENT TIMER PROTEINS AND METHODS  
; TITLE OF INVENTION: FOR THEIR USE  
; FILE REFERENCE: CLON-077CIP  
; CURRENT APPLICATION NUMBER: US/10/315,920  
; PRIOR APPLICATION NUMBER: 60/211,607  
; PRIOR FILING DATE: 2002-12-09  
; PRIOR FILING DATE: 2000-06-14  
; PRIOR APPLICATION NUMBER: PCT/US01/19097  
; PRIOR FILING DATE: 2001-06-13  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: variant of sequence from Discosoma sp.  
US-10-315-920-1

Alignment Scores:  
Pred. No.: 4.4e-143 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0  
DB: 16 Gaps: 0  
US-10-006-922A-12 (1-225) x US-10-315-920-1 (1-678)  
QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 1 ATGCGCTCCTCCAAGAACGTTCATCAAGGAGTTTCATCGCTTCAAGGTGCGCATGGAGGC 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluArgProTyrGluGly 40  
Db 61 ACCGTGAACGGCCACGAGTTCGAGATCGAGGCGAGGGCGAGGCCGCCCTACGAGGSC 120  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 121 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTGCCTTCGCCTGGGACATC 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 181 CTGTCCCCCAGTTCCAGTACGGCTCCAAGGTGTACGTGAAGCACCCCGCCGACATCCCC 240  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
Db 241 GACTACAAGAAAGCTGTCTTCCCAGAGGGCTTCAAGTGGGAGCGCGTGATGAACCTTCGAG 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGGCGCGTGGTGACCGTGACCCAGGACTCCTCCCTGCAGACGGCTGCTTCATCTAC 360  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 361 AAGGTGAAGTTCATCGCGTGAACCTTCCCTCCGACGGCCCGCTGATGCAGAAGAGACC 420  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 421 ATGGGCTGGGAGGCCTCCACCGAGCGCTGTACCCCGCGACGGCGTGCTGAAGGGCGAG 480  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 481 ATCCACAAGGCCCTGAAGCTGAAGACGGCGGCCACTACCTGTTGGAGTTTCAAGTCCATC 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 541 TACATGGCCAAAGAGCCCGTGCAGTCCCGGCTACTACTAGTGGACTCCAAGCTGGAC 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATCACCTCCACAAACGAGGACTACACCATCGTGGAGCAGTACGAGCGCACCGAGGGCGCG 660  
QY 221 HisHisLeuPheLeu 225  
Db 661 CACCACCTGTTTCCTG 675

RESULT 6

US-10-132-067-3  
; Sequence 3, Application US/10132067  
; Publication No. US20030203355A1  
; GENERAL INFORMATION:  
; APPLICANT: Bradbury, Andrew  
; APPLICANT: Zeytun, Ahmet  
; APPLICANT: Waldo, Geoffrey  
; APPLICANT: The Regents of the University of California  
; TITLE OF INVENTION: Fluorobodies: Binding Ligands with Intrinsic  
; TITLE OF INVENTION: Fluorescence  
; FILE REFERENCE: 021362-000600US  
; CURRENT APPLICATION NUMBER: US/10/132,067  
; CURRENT FILING DATE: 2002-04-24  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Discosoma sp.

; FEATURE:
; OTHER INFORMATION: red fluorescent protein (dsRED)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(678)
; OTHER INFORMATION: dsRED
US-10-132-067-3

Alignment Scores:
Pred. No.: 4.4e-143 Length: 678
Score: 1214.00 Matches: 225
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-10-006-922A-12.(1-225) x US-10-132-067-3 (1-678)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
Db 1 ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTCCGATGGAAGGA 60
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
Db 61 ACGGTCAATGGGCACGAGTTTGAATAGAGGCCGAAGAGAGGGAGGCCATACGAAGGC 120
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60
Db 121 CACAATACCGTAAAGCTTAAAGGTAAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80
Db 181 TTGTCACCACCAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA 240
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100
Db 241 GACTATAAAAGCTGTCTTCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAA 300
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
Db 301 GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTCAGGATGGCTGTTTCATCTAC 360
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
Db 361 AAGGTCAAGTTTCATTGGCGTGAACCTTTCTCCGATGGACCTGTTATGCAAAAGAGACA 420
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
Db 421 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCAATTACCTAGTTGAATTCAAAAGTATT 540
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrValAspSerLysLeuAsp 200
Db 541 TACATGGCAAAGAAGCCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
Db 601 ATAACAAGCCACAACGAAGACTATACAAATCGTTGAGCAGTATGAAAGAACCCGAGGACGC 660
QY 221 HisHisLeuPheLeu 225
Db 661 CACCATCTGTTCCCTT 675

RESULT 7

US-10-335-517-6
; Sequence 6, Application US/10335517
; Publication No. US20030207248A1
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: TSJEN, Roger

; APPLICANT: GONZALEZ, Jesus
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY OPTICAL METHODS
; FILE REFERENCE: REGEN1290-5
; CURRENT APPLICATION NUMBER: US/10/335,517
; CURRENT FILING DATE: 2002-12-31
; PRIOR APPLICATION NUMBER: US/09/967,772
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 09/459,956
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 08/765,860
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: PCT/ US96/09652
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: ~~US 08/481,977~~
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Discosoma sp "red"
US-10-335-517-6

Alignment Scores:
Pred. No.: 4.4e-143 Length: 678
Score: 1214.00 Matches: 225
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-335-517-6 (1-678)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
Db 1 ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTCCGATGGAAGGA 60
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
Db 61 ACGGTCAATGGGCACGAGTTTGAATAGAGGCCGAAGAGAGGGAGGCCATACGAAGGC 120
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60
Db 121 CACAATACCGTAAAGCTTAAAGGTAAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80
Db 181 TTGTCACCACCAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA 240
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100
Db 241 GACTATAAAAGCTGTCTTCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAA 300
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
Db 301 GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTCAGGATGGCTGTTTCATCTAC 360
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
Db 361 AAGGTCAAGTTTCATTGGCGTGAACCTTTCTCCGATGGACCTGTTATGCAAAAGAGACA 420
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
Db 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTATGGCGTGTGAAAGGAGAG 480
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
Db 481 ATTCATAAGGCTCTGAAGCTGAAAGACCGGTGGTTCATTACCTAGTTGAATTCAAAAGTATT 540
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrValAspSerLysLeuAsp 200
Db 541 TACATGGCAAAGAAGCCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600

```
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
Db 601 ATAACAAGCCACAAAGAACTATACAATCGTTGACAGTATGAAAGAACCGAGGGACGC 660

QY 221 HisHisLeuPheLeu 225
Db 661 CACCATCTGTTCCCTT 675

RESULT 8
US-10-334-288-6
; Sequence 6, Application US/10334288
; Publication No. US20040002123A1
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: TSJEN, Roger
; APPLICANT: GONZALEZ, Jesus
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY OPTICAL METHODS
; FILE REFERENCE: REGEN1290-5
; CURRENT APPLICATION NUMBER: US/10/334,288
; CURRENT FILING DATE: 2002-12-31
; PRIOR APPLICATION NUMBER: US/09/967,772
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 09/459,956
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 08/765,860
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: PCT/ US96/09652
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: US 08/481,977
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Discosoma sp "red"
US-10-334-288-6

Alignment Scores:
Pred. No.: 4.4e-143 Length: 678
Score: 1214.00 Matches: 225
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-334-288-6 (1-678)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
Db 1 ATGAGGTCTTCCAAGATGTTATCAAGGAGTTTCAGGTTTAAGGTTTCGATGGAAGGA 60

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
Db 61 ACGGTCAATGGGCACGAGTTTGAAATAGAACGCCAAGGAGGGAGGCCATACGAAGGC 120

QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60
Db 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180

QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80
Db 181 TTGTCAACCAAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAGGTCATGAACTTTGA 240

QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100
Db 241 GACTATAAAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAGGTCATGAACTTTGA 300

QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
Db 301 GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTCAGGATGGCTGTTTCATCTAC 360

QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
```

```
Db 361 AAGGTCAAGTTCAATGGCGTGAACCTTCCTTCGGATGGACCTGTATGCAAAAGAAGACA 420

QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
Db 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGATCCTCGTGATGGCGTGTGAAAGGAGAG 480

QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
Db 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT 540

QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrValAspSerLysLeuAsp 200
Db 541 TACATGGCAAAGAAGCCTGTGCAGCTACCAAGGTACTACTATGTGACTCCAAACTGGAT 600

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
Db 601 ATAACAAGCCACAAAGAACTATACAATCGTTGAGCAGTATGAAAGAACCAGGAGCGC 660

QY 221 HisHisLeuPheLeu 225
Db 661 CACCATCTGTTCCCTT 675

RESULT 9
US-10-311-030-5
; Sequence 5, Application US/10311030
; Publication No. US20040171107A1
; GENERAL INFORMATION:
; APPLICANT: Nelson, David
; APPLICANT: Zamiaara, Elize
; APPLICANT: Tsien, Roger
; TITLE OF INVENTION: MODIFIED FLUORESCENT PROTEINS
; FILE REFERENCE: 15916-032US1
; CURRENT APPLICATION NUMBER: US/10/311,030
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: PCT/US01/04625
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: US 60/184,732
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Discosoma sp.
US-10-311-030-5
```

```
Alignment Scores:
Pred. No.: 4.4e-143 Length: 678
Score: 1214.00 Matches: 225
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 19 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-311-030-5 (1-678)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
Db 1 ATGAGGTCTTCCAAGATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGATGGAAGGA 60

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
Db 61 ACGGTCAATGGGCACGAGTTTGAAATAGAACGCCAAGGAGGGAGGCCATACGAAGGC 120

QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60
Db 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180

QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80
Db 181 TTGTCAACCAAAATTTCAGTATGGAAGCAAGGTTATATGTCAAGCACCCCTGCGGACATACCA 240
```







US-10-006-922A-12 (1-225) x US-10-505-486-27 (1-678)

Qy	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
Db	1	ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTCATGAGGTTTAAAGTTTCGCATGGAAGGA	60
Qy	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
Db	61	ACGGTCAATGGGCACGAGTTTGAAATAGAAGCGAAGGAGAGGGGAGCCATACGAAGGC	120
Qy	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle	60
Db	121	CACAATACCGTAAAGCTTAAGGTAAACCAAGGGGGGACCTTTGCCATTTGCTGGGATATT	180
Qy	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	181	TTGTCAACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA	240
Qy	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
Db	241	GACTATATAAAAGCTGTCAATTTCTCGAAGGATTATAATGGGAAAGGGTCATGAACCTTTGAA	300
Qy	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	301	GACGGTGGCGTCGTTACTGTATACCCAGGATTCCAGTTTGCAGGATGGCTGTTTCATCTAC	360
Qy	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
Db	361	AAGGTCAAGTTCATTGGCGTGAACTTTCTTCCGATGGACCTGTTATGCAAAAAGAAGACA	420
Qy	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	421	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTATGGCGTGTGAAAGGAGAG	480
Qy	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db	481	ATTCAATAAGGCTCTGAAGCTGAAAGACCGGTGGTCATTACCTAGTTGAATTCAAAAAGTATT	540
Qy	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrValAspSerLysLeuAsp	200
Db	541	TACATGGCAAAAGAAGCCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT	600
Qy	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	601	ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAAACCGAGGGACGC	660
Qy	221	HisHisLeuPheLeu	225
Db	661	CACCATCTGTTCCCTT	675

RESULT 12

```

US-09-999-745-66
; Sequence 66, Application US/09999745
; Patent No. US20020157120A1
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: CIRCULARLY PERMUTED FLUORESCENT PROTEIN INDICATORS
; FILE REFERENCE: REGEN1470-1
; CURRENT APPLICATION NUMBER: US/09/999,745
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 09/316,920
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 66
; LENGTH: 859
; TYPE: DNA
; ORGANISM: Discosoma sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)..(731)

```

; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 11  
; LENGTH: 859  
; TYPE: DNA  
; ORGANISM: Discosoma sp.  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (54)..(731)  
US-09-866-538-11

Alignment Scores:  
Pred. No.: 6.12e-143 Length: 859  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-10-006-922A-12 (1-225) x US-09-866-538-11 (1-859)

QY	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
Db	54	ATGAGGTCCTCCAAGATGTTATCAAGGAGTTCATGAGGTTTAAGGTTCCGATGGAAGGA	113
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
Db	114	ACGGTCAATGGGCACGAGTTTGAAATAGAGGCGAAGGAGAGGGGAGGCATACGAAGGC	173
QY	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle	60
Db	174	CACAATACCGTAAAGCTTAAGGTAAACCAAGGGGGACCTTTGCCATTTGCTGGGATATT	233
QY	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	234	TTGTCAACCAAAAGCTGTCATTTCTGAAGGATTAAATGGGAAAGGTCATGAACCTTTGAA	293
QY	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
Db	294	GACTATAAAAGCTGTCATTTCTGAAGGATTAAATGGGAAAGGTCATGAACCTTTGAA	353
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	354	GACGGTGGCGTCGTACTGTAAACCCAGGATTCAGTTTCAGGATGGCTGTTTCATCTAC	413
QY	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
Db	414	AAGGTCAAGTTCATTGGCGTGAACTTTCCCTCCGATGGACCTGTTTATGCAAAAGAAGACA	473
QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	474	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG	533
QY	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db	534	ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT	593
QY	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	594	TACATGGCAAGAGCCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT	653
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	654	ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC	713
QY	221	HisHisLeuPheLeu	225
Db	714	CACCATCTGTTCCCTT	728

RESULT 14

US-09-794-308-11  
; Sequence 11, Application US/09794308  
; Publication No. US20030170911A1  
; GENERAL INFORMATION:

; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: TSJEN, Roger  
; APPLICANT: ZACHARIAS, David  
; APPLICANT: BAIRD, Geoffrey  
; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS  
; FILE REFERENCE: REGEN1530  
; CURRENT APPLICATION NUMBER: US/09/794,308  
; CURRENT FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 11  
; LENGTH: 859  
; TYPE: DNA  
; ORGANISM: Discosoma sp.  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (54)..(731)  
US-09-794-308-11

Alignment Scores:

Pred. No.: 6.12e-143 Length: 859  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-10-006-922A-12 (1-225) x US-09-794-308-11 (1-859)

QY	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
Db	54	ATGAGGTCCTCCAAGATGTTATCAAGGAGTTCATGAGGTTTAAGGTTCCGATGGAAGGA	113
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
Db	114	ACGGTCAATGGGCACGAGTTTGAAATAGAGGCGAAGGAGAGGGGAGGCATACGAAGGC	173
QY	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle	60
Db	174	CACAATACCGTAAAGCTTAAGGTAAACCAAGGGGGACCTTTGCCATTTGCTGGGATATT	233
QY	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	234	TTGTCAACCAAAATTTCAAGTATGGAAGCAAGGTTATATGTCAAGCACCTCCGACATACCA	293
QY	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
Db	294	GACTATAAAAGCTGTCATTTCTGAAGGATTAAATGGGAAAGGTCATGAACCTTTGAA	353
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	354	GACGGTGGCGTCGTACTGTAAACCCAGGATTCAGTTTCAGGATGGCTGTTTCATCTAC	413
QY	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
Db	414	AAGGTCAAGTTCATTGGCGTGAACTTTCCCTCCGATGGACCTGTTTATGCAAAAGAAGACA	473
QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	474	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG	533
QY	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db	534	ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT	593
QY	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	594	TACATGGCAAGAGCCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT	653
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	654	ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC	713

Qy 221 HisHisLeuPheLeu 225  
Db 714 CACCATCTGTTCTT 728

## RESULT 15

```

US-09-865-291-11
; Sequence 11, Application US/09865291
; Publication No. US20030186229A1
; GENERAL INFORMATION:
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: TSIEN, Roger
; APPLICANT: TING, Alice
; APPLICANT: ZHANG, Jin
; TITLE OF INVENTION: EMISSION RATIO-METRIC INDICATORS OF PHOSPHORYLATION
; FILE REFERENCE: REGEN1550
; CURRENT APPLICATION NUMBER: US/09/865,291
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 859
; TYPE: DNA
; ORGANISM: Discosoma sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)..(731)
;
US-09-865-291-11

```

Alignment Scores:		
pred. No.:	6.12e-143	859
Score:	1214.00	225
Percent Similarity:	100.00%	Length:
Best Local Similarity:	100.00%	Matches:
Query Match:	100.00%	Conservative:
DB:	10	Mismatches:
		Indels:
		Gaps:

US-10-006-922A-12 (1-225) x US-09-865-291-11 (1-859)

Qy	1	MetArgSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
Db	54	ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTTCATGAGGTTTAAAGTTCGCATGGAAGGA	113
Qy	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly	40
Db	114	ACGGTCAATGGGCACGAGTTTGAAATAGAAAGGCGAAGGAGAGGGGAGGCCATACGAAAGGC	173
Qy	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle	60
Db	174	CACAAATACCGTAAAGCTTAAGGTAACCAAGGGGGGACCTTTGCCATTTGCTTGGGATATT	233
Qy	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	234	TTGTCAACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA	293
Qy	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
Db	294	GACTATAAAAAGCTGTCTATTCCTGAAGGATTTAATGGGAAAGGGTCATGAACCTTTGAA	353
Qy	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	354	GACGGTGGCGTCGTTACTGTATACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC	413
Qy	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
Db	414	AAGGTCAAGTTTCATTGGCGTGAACTTTCCTTCCGATGGACCTGTATTGCAAAAAGAAGACA	473
Qy	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	474	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGTGAAGGAGAG	533
Qy	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db	534	ATTTCATAAGGCTCTGAAGCTGAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT	593

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 1, 2005, 18:23:08 ; Search time 186 Seconds  
(without alignments)  
1979.367 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMFKVRMEG.....EDYIVEQYERTEGRHHLFL 225

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/uspto spool/US10006922/runat 30062005 110617 19116/app query.fasta\_1.391  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10006922@cgn\_1 1 69 @runat 30062005 110617 19116 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq: \*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1214	100.0	678	3	US-09-459-956-6
2	1210	99.7	723	4	US-10-152-296-1
3	729.5	60.1	699	3	US-09-459-956-5
4	667	54.9	801	3	US-09-459-956-7
5	495	40.8	690	3	US-09-459-956-2
6	486.5	40.1	1079	3	US-09-609-161B-15
7	486.5	40.1	1079	4	US-09-626-581D-64
8	486.5	40.1	1079	4	US-09-415-765B-64
9	486.5	40.1	1079	4	US-09-626-580C-64
10	486.5	40.1	1085	3	US-09-277-716-15
11	485.5	40.0	696	3	US-09-459-956-4
12	485.5	40.0	720	4	US-09-839-650-1

13	477.5	39.3	696	3	US-09-459-956-3	Sequence 3, Appli
14	474.5	39.1	1021	4	US-09-839-650-2	Sequence 2, Appli
15	474.5	39.1	1104	3	US-09-277-716-30	Sequence 30, Appli
16	474.5	39.1	1104	3	US-09-609-161B-30	Sequence 30, Appli
17	474.5	39.1	1279	3	US-09-277-716-31	Sequence 31, Appli
18	474.5	39.1	1279	3	US-09-609-161B-31	Sequence 31, Appli
19	468.5	38.6	1482	4	US-09-977-897-1	Sequence 1, Appli
20	218.5	18.0	1559	3	US-09-049-475-6	Sequence 6, Appli
21	215.5	17.8	717	4	US-09-023-946B-3	Sequence 3, Appli
22	215.5	17.8	4196	4	US-09-453-313-1	Sequence 1, Appli
23	215.5	17.8	4199	3	US-09-204-117B-1	Sequence 1, Appli
24	215.5	17.8	6232	4	US-09-796-575-1	Sequence 1, Appli
25	215.5	17.8	7455	4	US-09-220-557-19	Sequence 19, Appli
26	215.5	17.8	7686	3	US-09-502-710-23	Sequence 23, Appli
27	215.5	17.8	7686	3	US-09-502-710-26	Sequence 26, Appli
28	215.5	17.8	7686	3	US-09-502-711-23	Sequence 23, Appli
29	215.5	17.8	7686	3	US-09-502-711-26	Sequence 26, Appli
30	215.5	17.8	7686	4	US-09-565-616A-2	Sequence 2, Appli
31	215.5	17.8	7687	3	US-09-502-710-24	Sequence 24, Appli
32	215.5	17.8	7687	3	US-09-502-711-24	Sequence 24, Appli
33	212.5	17.5	717	2	US-08-818-604-31	Sequence 31, Appli
34	212.5	17.5	717	4	US-09-346-946-31	Sequence 31, Appli
35	212.5	17.5	717	4	US-09-023-946B-21	Sequence 21, Appli
36	212.5	17.5	764	2	US-08-818-604-30	Sequence 30, Appli
37	212.5	17.5	764	3	US-08-819-612-21	Sequence 21, Appli
38	212.5	17.5	764	4	US-09-346-946-30	Sequence 30, Appli
39	212.5	17.5	764	4	US-09-872-364-21	Sequence 21, Appli
40	211.5	17.4	1665	2	US-08-771-850A-1	Sequence 1, Appli
41	210.5	17.3	716	1	US-08-337-915A-1	Sequence 1, Appli
42	210.5	17.3	716	1	US-08-753-143-1	Sequence 1, Appli
43	210.5	17.3	716	2	US-08-792-553-1	Sequence 1, Appli
44	210.5	17.3	716	3	US-08-753-144-1	Sequence 1, Appli
45	210.5	17.3	716	3	US-09-094-359-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-09-459-956-6  
; Sequence 6, Application US/09459956  
; Patent No. 6342379  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger Y.  
; APPLICANT: Gonzalez, III, Jesus E.  
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY  
; TITLE OF INVENTION: OPTICAL METHODS  
; FILE REFERENCE: REGEN1290-4  
; CURRENT APPLICATION NUMBER: US/09/459,956  
; CURRENT FILING DATE: 1999-12-13  
; PRIOR APPLICATION NUMBER: 08/765,860  
; PRIOR FILING DATE: 1999-05-08  
; PRIOR APPLICATION NUMBER: 08/481,977  
; PRIOR FILING DATE: 1995-06-07  
; PRIOR APPLICATION NUMBER: PCT/US96/09652  
; PRIOR FILING DATE: 1996-06-06  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Discosoma sp  
US-09-459-956-6

Alignment Scores:  
Pred. No.: 1.71e-160 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-10-006-922A-12 (1-225) x US-09-459-956-6 (1-678)



```
QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
Db 1 ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTGCGATGGAAGGA 60

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
Db 61 ACGGTCAATGGCAGCAGAGTTTGAAATAGAGCGCAAGGAGAGAGGGGAGGCCATACGAAGGC 120

QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60
Db 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTGCTTGGGATATT 180

QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80
Db 181 TTGTCAACCACAAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240

QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100
Db 241 GACTATAAAAAGTGTCTCATTTCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAA 300

QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
Db 301 GACGGTGGCGTCGTACTGTAAACCCAGGATTCCAGTTTGCAAGATGGCTGTTTCATCTAC 360

QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
Db 361 AAGGTCAAGTTTCATTTGGCGTGAACCTTTCCTTCCGATGGACCTGTTATGCAAAAAGAGACA 420

QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
Db 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480

QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
Db 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTGCATTACCTAGTTGAATTCAAAAGTATT 540

QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
Db 541 TACATGGCAAGAAGAGCCTGTGCAGCTACCAGGCTACTACTATGTTGACTCCAAACTGGAT 600

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
Db 601 ATAACAAGCCACACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCAGGGACGC 660

QY 221 HisHisLeuPheLeu 225
Db 661 CACCATCTGTTCCCTT 675
```

RESULT 2
US-10-152-296-1

```
; Sequence 1, Application US/10152296
; Patent No. 6723537
; GENERAL INFORMATION:
; APPLICANT: Peelle, Beau
; APPLICANT: Rigel Pharmaceuticals, Incorporated
; TITLE OF INVENTION: Directed Evolution of Protein in Mammalian Cells
; FILE REFERENCE: 021044-000110US
; CURRENT APPLICATION NUMBER: US/10/152,296
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/291,871
; PRIOR FILING DATE: 2001-05-18
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:mammalian
; OTHER INFORMATION: codon-optimized variant (DsRED) of Discosoma sp.
; OTHER INFORMATION: "red" red fluorescent protein (RFP)
; FEATURE:
; NAME/KEY: CDS
```

```
; LOCATION: (1)..(723)
; OTHER INFORMATION: DsRED
US-10-152-296-1
```

```
Alignment Scores:
Pred. No.: 6.93e-160 Length: 723
Score: 1210.00 Matches: 224
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.56% Mismatches: 0
Query Match: 99.67% Indels: 0
DB: 4 Gaps: 0
```

US-10-006-922A-12 (1-225) x US-10-152-296-1 (1-723)

```
QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
Db 4 GTGCGCTCCTCCAAGAACGTCATCAAGGAGTTTCATGCGCTTCAAGGTGCGCATGGAGGGC 63

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
Db 64 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGCCGCCCTACGAGGGC 123

QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60
Db 124 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTGCCCTTCGCCTGGGACATC 183

QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80
Db 184 CTGTCCCCCAGTTCCAGTACGGCTCCAAGGTGTACGTGAAGCACCCCGCCGACATCCCC 243

QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100
Db 244 GACTACAAGAAGCTGTCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACCTTCGAG 303

QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
Db 304 GACGGCGCGTGGTGACCGTGACCCAGGACTCCTCCCTGCAGACGGCTGCTTTCATCTAC 363

QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
Db 364 AAGGTGAAGTTCATCGGCGTGAACCTTCCCCCTCCGACGGCCCGCTAATGCAGAAGAAGACC 423

QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
Db 424 ATGGGCTGGGAGGCCCTCCACCGAGCGCTGTACCCCCCGACGGCGTGTGAAGGGCGAG 483

QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
Db 484 ATCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACCTGGTGGAGTTCAAGAGTATC 543

QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
Db 544 TACATGGCCAAGAAGCCCGTGCAGCTGCCCCGGCTACTACTACGTGGACTCCAAGCTGGAC 603

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
Db 604 ATCACCTCCCAACACGAGGACTACACCATCGTGGAGCAGTACGAGCGCACCGAGGGCCGC 663

QY 221 HisHisLeuPheLeu 225
Db 664 CACCACCTGTTCCCTG 678
```

```
RESULT 3
US-09-459-956-5
; Sequence 5, Application US/09459956
; Patent No. 6342379
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Gonzalez, III, Jesus E.
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY
; TITLE OF INVENTION: OPTICAL METHODS
; FILE REFERENCE: REGEN1290-4
; CURRENT APPLICATION NUMBER: US/09/459,956
```

```
; CURRENT FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: 08/765,860
; PRIOR FILING DATE: 1999-05-08
; PRIOR APPLICATION NUMBER: 08/481,977
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: PCT/US96/09652
; PRIOR FILING DATE: 1996-06-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 699
; TYPE: DNA
; ORGANISM: Discosoma striata
US-09-459-956-5
```

```
Alignment Scores:
Pred. No.: 1.95e-92 Length: 699
Score: 729.50 Matches: 133
Percent Similarity: 76.00% Conservative: 38
Best Local Similarity: 59.11% Mismatches: 53
Query Match: 60.09% Indels: 1
DB: 3 Gaps: 1
```

US-10-006-922A-12 (1-225) x US-09-459-956-5 (1-699)

```
QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
Db 1 ATGAGTTGTTCCAGAGTGTGATCAAGGAAGAAATGTTGATCGATCTTTCATCTGGAAGGA 60

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
Db 61 ACGTTCAATGGGCACACTTTGAATAAAAGGCAAGGAAAGGACAGCCCTAATGAAGGC 120

QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60
Db 121 ACCAATACCGTCACGCTCGAGGTTACCAAGGGTGGACCTCTGCCATTTGGTTGGCATATT 180

QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80
Db 181 TTGTGCCACAAATTCAGTATGGAAACAAGGCATTGTGCCACCCTGACACATACAT 240

QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100
Db 241 GATTATCTAAAGCTGTCAATTCGGAGGGATATACATGGGAACGGTCCATCGCACTTTGAA 300

QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
Db 301 GACGGTGGCTTGTGTTGTATCACCATGATATCATGTTTGACAGGCACTGTTCTACTAC 360

QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
Db 361 GACATCAAGTTCACTGGCTTGAACCTTTCCTCCAAATGGACCCGTTGTGCAGAAGACA 420

QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
Db 421 ACTGGCTGGGAACCGGCACTGAGCGTTTGTATCCTCGTGATGGTGTGTGATAGGAGAC 480

QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
Db 481 ATCCATCATGCTCTGACAGTTGAAGGAGGTGGTTCATACGCATGTGACATTTAAAACTGTT 540

QY 181 TyrMetAlaLysLys---ProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeu 199
Db 541 TACAGGGCCAAGAGCCGCTTGAAGATGCCAGGTTATCACTATGTTGACACCAACTG 600

QY 200 AspIleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGly 219
Db 601 GTTATATGGAACAACGCAAGAAGATTTCATGAAAGTTGAGGAGCATGAAATCGCCGTTGCA 660

QY 220 ArgHisHisLeuPhe 224
Db 661 CGCCACCATCCGTTTC 675
```

```
RESULT 4
US-09-459-956-7
; Sequence 7, Application US/09459956
; Patent No. 6342379
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Gonzalez, III, Jesus E.
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY
; TITLE OF INVENTION: OPTICAL METHODS
; FILE REFERENCE: REGEN1290-4
; CURRENT APPLICATION NUMBER: US/09/459,956
; CURRENT FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: 08/765,860
; PRIOR FILING DATE: 1999-05-08
; PRIOR APPLICATION NUMBER: 08/481,977
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: PCT/US96/09652
; PRIOR FILING DATE: 1996-06-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 801
; TYPE: DNA
; ORGANISM: Clavularia sp
US-09-459-956-7
```

```
Alignment Scores:
Pred. No.: 1.45e-83 Length: 801
Score: 667.00 Matches: 121
Percent Similarity: 73.76% Conservative: 42
Best Local Similarity: 54.75% Mismatches: 58
Query Match: 54.94% Indels: 0
DB: 3 Gaps: 0
```

US-10-006-922A-12 (1-225) x US-09-459-956-7 (1-801)

```
QY 3 SerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGlyThrVal 22
Db 121 ACCACCATGGGTGTGATTAACCCAGACATGAAGATTAAAGCTGAAGATGGAAGAAATGTA 180

QY 23 AsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGlyHisAsn 42
Db 181 AACGGGCATGCTTTGTGATCGAAGGAGGAGGAGGAGAAAGCCTTACGATGGGACACAC 240

QY 43 ThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIleLeuSer 62
Db 241 ACTTTAAACCTGGAAGTGAAGGAAGGTGGCCTCTGCCTTTTCTTACGATATCTTGTCA 300

QY 63 ProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIleProAspTyr 82
Db 301 AACGCGTTCAGTACGGAAACAGAGCATTTGACAAAATACCCAGACGATATAGCAGACTAT 360

QY 83 LysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGluAspGly 102
Db 361 TTCAAGCAGTCGTTTCCCGAGGGATATTCTCTGGGAAAGAACCATGACTTTTGAAGACAAA 420

QY 103 GlyValValThrValThrGlnAspSerSerSerLeuGlnAspGlyCysPheIleTyrLysVal 122
Db 421 GGCATTGTCAAAGTGAAAGTAGACATAAGCATAGCATGGAGGAAGACTCCTTTTATCTATGAJATT 480

QY 123 LysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThrMetGly 142
Db 481 CGTTTTGATGGGATGAACCTTTCTCTCCCAATGGTCCGGTTATGCAGAAAAAACTTTGAAG 540

QY 143 TrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGluIleHis 162
Db 541 TGGGAACCATCCACTGAGATTATGTACGTGCGTGATGGAGTGTGTCGGAGATATTAGC 600

QY 163 LysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIleTyrMet 182
Db 601 CATTCCTCTGTTGCTGGAGGGAGGTGGCCATTACCGATGTGACTTCAAAAGTATTATACAAA 660

QY 183 AlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAspIleThr 202
```

Db 661 GCAAAAAAGTTGTCAAATGCCAGACTATCACTTTGTGGACCATCGCATTGAGATCTTG 720

QY 203 SerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHisHis 222

Db 721 AACCATGACAAGGATTACAACAAGTAACGCTGTATGAGAAATGCAGTTGCTCGCTATTCT 780

QY 223 Leu 223

Db 781 TTG 783

RESULT 5

US-09-459-956-2

; Sequence 2, Application US/09459956

; Patent No. 6342379

; GENERAL INFORMATION:

; APPLICANT: Tsien, Roger Y.

; APPLICANT: Gonzalez, Iii, Jesus E.

; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY

; TITLE OF INVENTION: OPTICAL METHODS

; FILE REFERENCE: REGEN1290-4

; CURRENT APPLICATION NUMBER: US/09/459,956

; CURRENT FILING DATE: 1999-12-13

; PRIOR APPLICATION NUMBER: 08/765,860

; PRIOR FILING DATE: 1999-05-08

; PRIOR APPLICATION NUMBER: 08/481,977

; PRIOR FILING DATE: 1995-06-07

; PRIOR APPLICATION NUMBER: PCT/US96/09652

; PRIOR FILING DATE: 1996-06-06

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 690

; TYPE: DNA

; ORGANISM: Anemonia majano

US-09-459-956-2

Alignment Scores:

Pred. No.: 1.62e-59 Length: 690

Score: 495.00 Matches: 92

Percent Similarity: 64.68% Conservative: 38

Best Local Similarity: 45.77% Mismatches: 69

Query Match: 40.77% Indels: 2

DB: 3 Gaps: 1

US-10-006-922A-12 (1-225) x US-09-459-956-2 (1-690)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20

Db 1 ATGGCTCTTTCAAACAAGTTTATCGGAGATGACATGAAAAATGACCTACCATATGGATGC 60

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40

Db 61 TGTGTCAATGGGCATTACTTTACCGTCAAAGGTGAAGGCAACGGGAAGCCATACGAAGG 120

QY 41 HisAsnThrValLysLeuLysValThr-----LysGlyGlyProLeuProPheAlaTrp 58

Db 121 ACGCAGACTTCGACTTTTAAAGTCACCATGGCCAACGGTGGGCCCTTGCATTCTCCTTT 180

QY 59 AspIleLeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAsp 78

Db 181 GACATACTATCTACAGTGTTCAAATATGGAATCGATGCTTTACTGCGTATCCTACCAGT 240

QY 79 IleProAspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsn 98

Db 241 ATGCCCGACTATTTCAAACAAGCATTTCCTGACGGAATGTCATATGAAAGGACTTTTACC 300

QY 99 PheGluAspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPhe 118

Db 301 TATGAAGATGGAGGATTGCTACAGCCAGTTGGGAAAATAAGCCTTAAGGCAACTGCTTT 360

QY 119 IleTyrLysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLys 138

Db 361 GAGCACAAATCCACGTTTTCATGGAGTGAACTTTCTGTGTATGGACCTGTGTATGGCGAAG 420

QY 139 LysThrMetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLys 158

Db 421 AAGACAACACTGGTGGACCCATCTTTTGAGAAAATGACTGTCTGCGATGGAATATTGAAG 480

QY 159 GlyGluIleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLys 178

Db 481 GGTGATGTCACCGCGTTCTCTCATGTGCAAGGAGGTGGCAATTACAGATGCCAATTCCAC 540

QY 179 SerIleTyrMetAlaLysLysProValGlnLeuProGlyTyrTyrValAspSerLys 198

Db 541 ACTTCTTACAAGACAAAAAACCGGTGACGATGCCCAAAACCATGTGGTGGAAACATCGC 600

QY 199 Leu 199

Db 601 ATT 603

RESULT 6

US-09-609-161B-15

; Sequence 15, Application US/09609161B

; Patent No. 6436682

; GENERAL INFORMATION:

; APPLICANT: Bryan, Bruce

; APPLICANT: Szent-Gyorgyi, Christopher

; APPLICANT: PROLUME, LTD.

; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE LUC:

; TITLE OF INVENTION: AND FLUORESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIGH

; TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS

; FILE REFERENCE: 24729-121B

; CURRENT APPLICATION NUMBER: US/09/609,161B

; CURRENT FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: 09/277,716

; PRIOR FILING DATE: 1999-03-26

; PRIOR APPLICATION NUMBER: 60/102,939

; PRIOR FILING DATE: 1998-10-01

; PRIOR APPLICATION NUMBER: 60/089,367

; PRIOR FILING DATE: 1998-06-15

; PRIOR APPLICATION NUMBER: 60/079,624

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 15

; LENGTH: 1079

; TYPE: DNA

; ORGANISM: Renilla mulleri

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (259)..(975)

; OTHER INFORMATION: Renilla mulleri Green Fluorescent Protein (GFP)

US-09-609-161B-15

Alignment Scores:

Pred. No.: 5.18e-58 Length: 1079

Score: 486.50 Matches: 97

Percent Similarity: 64.09% Conservative: 44

Best Local Similarity: 44.09% Mismatches: 72

Query Match: 40.07% Indels: 7

DB: 3 Gaps: 3

US-10-006-922A-12 (1-225) x US-09-609-161B-15 (1-1079)

QY 2 ArgSerSerLysAsnValIleLys-----GluPheMetArgPheLysVal 16

Db 256 AAGATGAGTAAACAAATATTGAAGAACACTTGTTTACAAGAAGTAATGTCGTATAAAGTA 315

QY 17 ArgMetGluGlyThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArg 36

Db 316 AATCTGGAAGGAATTGTAAACAACACCATGTTTTTACAATGGAGGGTTCGGGCAAGGGAAT 375

QY 37 ProTyrGluGlyHisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPhe 56

Db 376 ATTTTATTCGGCAATCAACTGGTTCAGATTCTGTGTACGAAAGGGGCCCCCTGCTTTT 435



QY 57 AlaTrpAspIleLeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisPro 76  
Db 436 GCATTGTGATATTGTGTACACGAGCTTTTCAATATGCGAACCGTACTTTTACACGAATATCCG 495

QY 77 AlaAspIleProAspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgVal 96  
Db 496 AATGATATATCAGATTATTTTATACAATCAATTCAGCAGGATTTATGTATGAACGAACA 555

QY 97 MetAsnPheGluAspGlyGlyValValThrValThrGlnAspSerSerLysValTyrValLysHisPro 116  
Db 556 TTACGTTACGAAGATGCGGACTTTGTGAATTCGTTTCAGATATAAAATTTAATAGAAGAC 615

QY 117 CysPheIleTyrLysValLysPheIleGlyValAlaSerThrGluArgLeuTyrProArgAspGlyVal 136  
Db 616 AAGTTCGTCACAGAGTGAATACAAAGGTAGTAACTTCCAGATGATGGTCCCGTCATG 675

QY 137 GlnLysLysThrMetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyVal 156  
Db 676 CAGAAGACTATCTTAGGAATAGAGCCTTCATTTGAAGCCATGTACATGAATAATGGCGTC 735

QY 157 LeuLysGlyGluIleHisLysAlaLeuLysLysAspGlyGlyHisTyrLeuValGlu 176  
Db 736 TTGGTCGGCGAAGTAATCTTGTCTATAAACTAACTCTGGGAATATTATTATCATGTAC 795

QY 177 PheLysSerIleTyrMetAlaLysLysProVal---GlnLeuProGlyTyrTyrVal 195  
Db 796 ATGAAACATTAATGAAGTCGAAAGGTAGTAAAGGAGTTTCCTTCGTATCATTTTATT 855

QY 196 AspSerLysLeuAspIleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGlu 215  
Db 856 CAACATCGTTTGGAAAAGACT--TACGTAGAAGACGGGGGTTTCGTTGAACAGCATGAG 912

RESULT 7

US-09-626-581D-64  
; Sequence 64, Application US/09626581D  
; Patent No. 6548249  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, David  
; TITLE OF INVENTION: Fusions of Scaffold Proteins with Random Peptide  
; TITLE OF INVENTION: Libraries  
; FILE REFERENCE: A-66900-3/RMS  
; CURRENT APPLICATION NUMBER: US/09/626,581D  
; CURRENT FILING DATE: 2000-07-27  
; PRIOR APPLICATION NUMBER: 09/169,015  
; PRIOR FILING DATE: 1998-10-08  
; PRIOR APPLICATION NUMBER: 09/415,765  
; PRIOR FILING DATE: 1999-10-08  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 64  
; LENGTH: 1079  
; TYPE: DNA  
; ORGANISM: Renilla muelleri  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (259)..(975)  
; OTHER INFORMATION:  
US-09-626-581D-64

Alignment Scores:

Pred. No.: 5.18e-58 Length: 1079  
Score: 486.50 Matches: 97  
Percent Similarity: 64.09% Conservative: 44  
Best Local Similarity: 44.09% Mismatches: 72  
Query Match: 40.07% Indels: 7  
DB: 4 Gaps: 3

US-10-006-922A-12 (1-225) x US-09-626-581D-64 (1-1079)

QY 2 ArgSerSerLysAsnValIleLys-----GluPheMetArgPheLysVal 16  
Db 256 AAGATGAGTAAACAATATTGAAGAACACTTGTGTACAGAAGTAAATGTCTGTATAAGTA 315

QY 17 ArgMetGluGlyThrValAsnGlyHisGluPheGluGlyGluGlyGluGlyArg 36  
Db 316 AATCTGGAAGGAATTTGTAAACAACCATGTTTTCACATGGAGGGTTGCGCAAGGGAAT 375

QY 37 ProTyrGluGlyHisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPhe 56  
Db 376 ATTTTATTCGCAATCAACTGGTTTCAGATTCGTGTACGAAAGGGGCCACTGCCTTTT 435

QY 57 AlaTrpAspIleLeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisPro 76  
Db 436 GCATTGTGATATTGTGTACACGACTTTTCAATATGGAACCGTACTTTTACGAAATATCCG 495

QY 77 AlaAspIleProAspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgVal 96  
Db 496 AATGATATATCAGATTATTTTATACAATCAATTCAGATTTTATGTATGAACGAACA 555

QY 97 MetAsnPheGluAspGlyGlyValValThrValThrGlnAspSerSerLysValGlnAspGly 116  
Db 556 TTACGTTACGAAGATGGCGGACTTTGTGAATTCGTTTCAGATATAAATTAATAGAAGAC 615

QY 117 CysPheIleTyrLysValLysPheIleGlyValAlaSerThrGluArgLeuTyrProArgAspGlyVal 136  
Db 616 AAGTTCGTCACAGAGTGAATACAAAGGTAGTAACTTCCAGATGATGGTCCCGTCATG 675

QY 137 GlnLysLysThrMetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyVal 156  
Db 676 CAGAAGACTATCTTAGGAATAGAGCCTTCATTTGAAGCCATGTACATGAATAATGGCGTC 735

QY 157 LeuLysGlyGluIleHisLysAlaLeuLysLysAspGlyGlyHisTyrLeuValGlu 176  
Db 736 TTGGTCGGCGAAGTAATCTTGTCTATAAACTAACTCTGGGAATATTATTATCATGTAC 795

QY 177 PheLysSerIleTyrMetAlaLysLysProVal---GlnLeuProGlyTyrTyrVal 195  
Db 796 ATGAAACATTAATGAAGTCGAAAGGTAGTAAAGGAGTTTCCTTCGTATCATTTTATT 855

QY 196 AspSerLysLeuAspIleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGlu 215  
Db 856 CAACATCGTTTGGAAAAGACT--TACGTAGAAGACGGGGGTTTCGTTGAACAGCATGAG 912

RESULT 8

US-09-415-765B-64  
; Sequence 64, Application US/09415765B  
; Patent No. 6548632  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, David  
; TITLE OF INVENTION: Fusions of Scaffold Proteins with Random Peptide  
; TITLE OF INVENTION: Libraries  
; FILE REFERENCE: A66900-1/RMS/AMS  
; CURRENT APPLICATION NUMBER: US/09/415,765B  
; CURRENT FILING DATE: 1999-10-08  
; PRIOR APPLICATION NUMBER: 09/169,015  
; PRIOR FILING DATE: 1998-10-08  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 64  
; LENGTH: 1079  
; TYPE: DNA  
; ORGANISM: Renilla muelleri  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (259)..(975)  
; OTHER INFORMATION:  
US-09-415-765B-64

Alignment Scores:

Pred. No.: 5.18e-58 Length: 1079  
Score: 486.50 Matches: 97  
Percent Similarity: 64.09% Conservative: 44  
Best Local Similarity: 44.09% Mismatches: 72  
Query Match: 40.07% Indels: 7  
DB: 4 Gaps: 3





; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 15

; LENGTH: 1079

; TYPE: DNA

; ORGANISM: Renilla mulleri

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (259)..(975)

; FEATURE:

; OTHER INFORMATION: Renilla mulleri Green Fluorescent Protein (GFP)

US-09-277-716-15

Alignment Scores:

Pred. No.: 5.23e-58 Length: 1085  
Score: 486.50 Matches: 97  
Percent Similarity: 64.09% Conservative: 44  
Best Local Similarity: 44.09% Mismatches: 72  
Query Match: 40.07% Indels: 7  
DB: 3 Gaps: 3

US-10-006-922A-12 (1-225) x US-09-277-716-15 (1-1085)

QY 2 ArgSerSerLysAsnValIleLys-----GlupheMetArgPheLysVal 16

Db 256 AAGATGAGTAAACAAATATTGAAGAACACTTGTGTACAAGAAGTAATGTCGTATAAAGTA 315

QY 17 ArgMetGluGlyThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArg 36

Db 316 AATCTGGAAGGAATTGTAAACAACCATGTTTACAATGGAGGTTGCGGCAAGGAAT 375

QY 37 ProTyrGluGlyHisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProphe 56

Db 376 ATTTTATTCGGCAATCAACTGGTTCAGATTCTGTACGAAAGGGCCCCCACTGCCTTTT 435

QY 57 AlaTrpAspIleLeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisPro 76

Db 436 GCATTGTGATATTGTGTACCAGCTTTTCAATATGGCAACCGTACTTTTCAACCAATATCCG 495

QY 77 AlaAspIleProAspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgVal 96

Db 496 AATGATATATCAGATTATTATTATACAATCATTTCCAGCAGGATTATGATGAACGAACA 555

QY 97 MetAsnPheGluAspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGly 116

Db 556 TTACGTTACGAAGATGCGGACTTGTGAAATTCGTTCCAGATATAAATTAATAGAAGAC 615

QY 117 CysPheIleTyrLysValLysPheIleGlyValAsnPheProSerAspGlyProValMet 136

Db 616 AAGTTCGTCTACAGAGTGAATACAAAGGTAGTAACCTCCAGATGATGGTCCCGTCATG 675

QY 137 GlnLysLysThrMetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyVal 156

Db 676 CAGAAGACTATCTTAGGAATAGAGCCTTCATTGGAAGCCATGTACATGAATAATGGCGTC 735

QY 157 LeuLysGlyGluIleHisLysAlaLeuLysLeuLysAspGlyHisTyrLeuValGlu 176

Db 736 TTGGTCGGCGAAGTAATCTTGTCTATAAACTAAACTCTGGGAAATATTATTCATGTCCAC 795

QY 177 PheLysSerIleTyrMetAlaLysLysProVal---GlnLeuProGlyTyrTyrVal 195

Db 796 ATGAAAACATTAAATGAAGTCGAAAGGTGTAGTAAAGGAGTTTCCTTCGTATCATTTTATT 855

QY 196 AspSerLysLeuAspIleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGlu 215

Db 856 CAACATCGTTTGGAAAAGACT--TACGTAGAGACGGGGGTTTCGTTGAACAGCATGAG 912

RESULT 11

US-09-459-956-4

; Sequence 4, Application US/09459956

; Patent No. 6342379

; GENERAL INFORMATION:

; APPLICANT: Tsien, Roger Y.

; APPLICANT: Gonzalez, III, Jesus E.  
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY  
; FILE REFERENCE: REGEN1290-4

; CURRENT APPLICATION NUMBER: US/09/459,956

; CURRENT FILING DATE: 1999-12-13

; PRIOR APPLICATION NUMBER: 08/765,860

; PRIOR FILING DATE: 1999-05-08

; PRIOR APPLICATION NUMBER: 08/481,977

; PRIOR FILING DATE: 1995-06-07

; PRIOR APPLICATION NUMBER: PCT/US96/09652

; PRIOR FILING DATE: 1996-06-06

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 696

; TYPE: DNA

; ORGANISM: Zoanthus sp

US-09-459-956-4

Alignment Scores:

Pred. No.: 3.54e-58 Length: 696  
Score: 485.50 Matches: 95  
Percent Similarity: 64.71% Conservative: 37  
Best Local Similarity: 46.57% Mismatches: 67  
Query Match: 39.99% Indels: 5  
DB: 3 Gaps: 3

US-10-006-922A-12 (1-225) x US-09-459-956-4 (1-696)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20

Db 1 ATGGCTCATTCAAAGCACGGTCTAAAGAAGAAATGACAATGAAATACCATGGAAGGG 60

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40

Db 61 TCGGTCAACGACATATAATTTGTGATCAGCGGCGAAGGCATTGGATATCCGTTCAAAGGG 120

QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60

Db 121 AAACAGACTATTAACTGTGTGTGATCGAAGGGGACCATTGCCATTTTCCGAAGACATA 180

QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80

Db 181 TTGTCAGCTGGCTTTAAGTACGGAGACAGGATTTTCACTGAATATCCTCAAGACATAGTA 240

QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100

Db 241 GACTATTTCAGAAACTCGTGTCTGCTGGATATACATGGGCGAGGCTCTTTCTCTTTGAG 300

QY 101 AspGlyGlyValValThrValThrGlnAsp-----SerSerLeuGlnAspGlyCysPhe 118

Db 301 GATGGAGCAGTCTGCATATGCAATGTAGATATAACAGTGAGTGTCAAAGAAAAAAGTGCATT 360

QY 119 IleTyrLysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLys 138

Db 361 TATCATAGAGCATATTTAATGGAATGAATTTTCTGCTGATGGACCTGTGATGAAAAG 420

QY 139 LysThrMetGlyTrpGluAlaSerThrGluArgLeu-----TyrProArgAspGlyVal 156

Db 421 ATGACAACTAACTGGGAAGCATCTCTGCGAAGAGATCATGCCAGTACCTTAAGCAGGGGATA 480

QY 157 LeuLysGlyGluIleHisLysAlaLeuLysLeuLysAspGlyHisTyrLeuValGlu 176

Db 481 CTGAAAGGGGATGTCTCCATGTACCTCTTCTGAAGGATGGTGGCGGTGTACCGGTGCCAG 540

QY 177 PheLysSerIleTyrMetAlaLysLys---ProValGlnLeuProGlyTyrTyrVal 195

Db 541 TTCGACACAGTTTACAAAGCAAAAGTCTGTGCCAAAGTAAGATGCCGGAGTGCACTTCATC 600

QY 196 AspSerLysLeu 199

Db 601 CAGCATAAGCTC 612

```
RESULT 12
US-09-839-650-1
; Sequence 1, Application US/09839650
; Patent No. 6645761
; GENERAL INFORMATION:
; APPLICANT: Stratagene
; TITLE OF INVENTION: Humanized Polynucleotide Sequence Encoding Renilla Mulleri Green
; Patent No. 6645761
; TITLE OF INVENTION: Fluorescent Protein
; FILE REFERENCE: 25436/1755
; CURRENT APPLICATION NUMBER: US/09/839,650
; CURRENT FILING DATE: 2001-04-19
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanized R. mulleri polynucleotide
; NAME/KEY: misc_feature
; LOCATION: (1)..(720)
; OTHER INFORMATION: Humanized DNA sequence
US-09-839-650-1

Alignment Scores:
Pred. No.:      3.74e-58      Length:      720
Score:          485.50      Matches:      97
Percent Similarity: 64.22%      Conservative: 43
Best Local Similarity: 44.50%      Mismatches:  71
Query Match:     39.99%      Indels:       7
DB:              4          Gaps:        3

US-10-006-922A-12 (1-225) x US-09-839-650-1 (1-720)
QY      4 SerLysAsnValIleLys-----GluPheMetArgPheLysValArgMet 18
Db      7 AGCAAGCAGATCCTGAAGAACACCTGCCTGCAGGAGGTGATGAGCTACAAGGTGAACCTG 66
QY      19 GluGlyThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyr 38
Db      67 GAGGGCATCGTGAACAACACACGTGTTCAACCATGGAGGGCTCGGCAAGGCAACATCCTG 126
QY      39 GluGlyHisAsnThrValLysLeuLysValThrLysGlyGlyGlyProLeuPropheAlaTrp 58
Db      127 TTCGGCAACCAGCTGGTGCAGATCCCGCTGACCAAGGGCGCCCCCTTGCCTTC 186
QY      59 AspIleLeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAsp 78
Db      187 GACATCGTGAGCCCGCCTTCCAGTACGGCAACCGCACCTTCAACCAAGTACCCCAACGAC 246
QY      79 IleProAspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsn 98
Db      247 ATCAGCGACTACTTTCATCCAGAGCTTCCCGCGCGCTTCATGTACGAGCGCACCTGCGC 306
QY      99 PheGluAspGlyGlyValValThrValThrValThrGlnAspSerSerLeuGlnAspGlyCysPhe 118
Db      307 TACGAGGACGGCGCCTGGTGGAGATCCGCGAGCGACATCAACCTGATCGAGGACAAAGTTC 366
QY      119 IleTyrLysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLys 138
Db      367 GTGTACCGCGTGGAGTACAAGGCGAGCAACTTCCCGACGACGCGCCCGTGTATGCAGAAG 426
QY      139 LysThrMetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLys 158
Db      427 ACCATCCTGGGCATCGAGCCCAGCTTCGAGGCCATGTACATGAACAACGCGGTGCTGGTG 486
QY      159 GlyGluIleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLys 178
Db      487 GCGAGGTGATCCTGGTGTACAAGCTGAACAGCGGCAAGTACTACAGTGCACATGAAG 546
QY      179 SerIleTyrMetAlaLysLysProVal---GlnLeuProGlyTyrTyrTyrValAspSer 197
```

```
Db      547 ACCCTGATGAAGACGAAGGGCGTGGTGAAGGAGTTCCCTCTTACCACCTTCATCCAGCAC 606
QY      198 LysLeuAspIleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGlu 215
Db      607 CGCCTGGAGAAGACC--TACGTGGAGGACGGCGGCTTCGTGGAGCAGCAGCAG 657

RESULT 13
US-09-459-956-3
; Sequence 3, Application US/09459956
; Patent No. 6342379
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Gonzalez, III, Jesus E.
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY
; TITLE OF INVENTION: OPTICAL METHODS
; FILE REFERENCE: REGEN1290-4
; CURRENT APPLICATION NUMBER: US/09/459,956
; CURRENT FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: 08/765,860
; PRIOR FILING DATE: 1999-05-08
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: PCT/US96/09652
; PRIOR FILING DATE: 1996-06-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 696
; TYPE: DNA
; ORGANISM: Zoanthus sp
US-09-459-956-3
```

```
Alignment Scores:
Pred. No.:      4.7e-57      Length:      696
Score:          477.50      Matches:      93
Percent Similarity: 64.22%      Conservative: 38
Best Local Similarity: 45.59%      Mismatches:  68
Query Match:     39.33%      Indels:       5
DB:              3          Gaps:        3

US-10-006-922A-12 (1-225) x US-09-459-956-3 (1-696)
QY      1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
Db      1 ATGGCTCAGTCAAAGCACGGTCTTAACAAAAAGAAATGACAAATGAAATACCGTATGGAAGG 60
QY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
Db      61 TGCCTCATGGACATAAATTTGTGATCACGGGAGGGCATTGGATATCCGTTCAAAGG 120
QY      41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuPropheAlaTrpAspIle 60
Db      121 AAACAGGCTATTAACTCTGTGTGTGGTGAAGGTGGACCATTGCCATTGCCGAAGACATA 180
QY      61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80
Db      181 TTGTCAGCTGCCTTTAACTACCGAAACACAGGGTTTTCACCTGAATATCCTCAAGACATAGTT 240
QY      81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100
Db      241 GACTATTCAAGAACTCGTGTCTCTGTGGATATACATGGGACAGGTCTTTTCTCTTTGAG 300
QY      101 AspGlyGlyValValThrValThrValThrGlnAsp-----SerSerLeuGlnAspGlyCysPhe 118
Db      301 GATGGAGCAGTTTGCATATGTATATGCAGATATAAACAGTGAAGTGTGAAGAAAACTGCATG 360
QY      119 IleTyrLysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLys 138
Db      361 TATCATGAGTCCAAATTTTATGGAGTGAATTTTCCTGCTGATGACACCTGTGATGAAAAAG 420
QY      139 LysThrMetGlyTrpGluAlaSerThrGluArgLeu-----TyrProArgAspGlyVal 156
```



Db	421	ATGACAGATAAACTGGGAGCCATCTTCGGAGAAGATCATACCAGTACCTTAAGCAGGGGATA	480
QY	157	LeuLysGlyGluIleHisLysAlaLeuLysLeuLysAspGlyGlyHistyLeuValGlu	176
Db	481	TTGAAGGGGATGTCCTCCATGTACCTCCTCTGAAGGATGGTGGCGTTTACGGTGCCAA	540
QY	177	PheLysSerIleTyrMetAlaLysLys--ProValGlnLeuProGlyTyrTyrTyrVal	195
Db	541	TTCGACACAGTTTACAAAGCAAAGTCTGTGCCAAGAAAGATGCCGGACTGGCACTTCATC	600
QY	196	AspSerLysLeu	199
Db	601	CAGCATAAGCTC	612

```

RESULT 14
US-09-839-650-2
; Sequence 2, Application US/09839650
; Patent No. 6645761
; GENERAL INFORMATION:
; APPLICANT: Stratagene
; TITLE OF INVENTION: Humanized Polynucleotide Sequence Encoding Renilla Mulleri Green
; Patent No. 6645761
; TITLE OF INVENTION: Fluorescent Protein
; FILE REFERENCE: 25436/1755
; CURRENT APPLICATION NUMBER: US/09/839,650
; CURRENT FILING DATE: 2001-04-19
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1021
; TYPE: DNA
; ORGANISM: Renilla muelleri
; FEATURE:
; NAME/KEY: exon
; LOCATION: (259)..(976)
US-09-839-650-2

```

US-10-006-922A-12 (1-225) x US-09-839-650-2 (1-1021)	
QY	2 ArgSerSerLysAsnValIleLys-----GlupHeMetArgPheLysVal 16
	:::       :::::         :::::
Db	256 AAGATGAGTAAACAAATATTGAAGAACACTTGTTTTACAAGAAAGTAATGTCGTATATAAGTA 315
QY	17 ArgMetGluGlyThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArg 36
	:::                         :::::       :::::
Db	316 AATCTGGAAGGAATTGTAAACAACCATGTTTTTACAATGGAGGGTTCGCGCAAAACCGAAT 375
QY	37 ProTyrGluGlyHisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProphe 56
	:::     :::::
Db	376 ATTTTATTCGGCAATCAACTGGITCACATTCGTGTACGAAAGGGGGCCCACTGCCTTTT 435
QY	57 AlaTrpAspIleLeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisPro 76
	:::       :::::                   :::::
Db	436 GCATTGTGATATTGTGTACCAGCTTTTCAATATGGCAACCGTACTTTTCACGAAATATCCG 495
QY	77 AlaAspIleProAspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgVal 96
	:::::
Db	496 AATGATATATCAGATTATTTTATACAAATCATTTCCAGCAGGATTTATGTATGAACGAACA 555
QY	97 MetAsnPheGluAspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGly 116
	::: :::::             :::::       :::::       :::::
Db	556 TTACGTTACGAAGATGGCGGACTTGTGAAATTCGTTTCAGATATAAAATTTTAATAGAAGAC 615
QY	117 CysPheIleTyrLysValLysPheIleGlyValAsnPheProSerAspGlyProValMet 136
	:::     :::::

Db 616 AAGTTCGTCTACAGAGTGGAATACAAAAGGTAGTAACCTCCACAGATGATGGTCCCGTCATG 675

Qy 137 GlnLysLysThrMetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyVal 156  
||||| :||| ||| ||| :||| :||| :||| :||| :|||

Db 676 CAGAAGACTATCTTAGGAATAGAGCCTTCATTTGAAGCCATGTACATGAATAATGSGCTC 735

Qy 157 LeuLysGlyGluIleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGlu 176  
||| ||||| :||| ||||| ||||| :||| :||| :|||

Db 736 TTGGTCGCGGAAGTAATCTTTGTCTATAAACTAACTCTGGGAAATATTATTTCATGTCAC 795

Qy 177 PheLysSerIleTyrMetAlaLysLysProVal---GlnLeuProGlyTyrTyrTyrVal 195  
||| :||| :||| ||| :||| :||| :||| :||| :|||

Db 796 ATGAAAACATAATGAAGTCGAAAGGTGTAGTAAAGGAGTTTCCTTCGTATCATTTATT 855

Qy 196 AspSerLysLeuAspIleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGlu 215  
:||| :||| :||| :||| :||| :||| :||| :||| :|||

Db 856 CAACATCGTTGGAAAAGACT--TACGTAGNAGACGGGGGTTTCGTTGAACAGCATGAG 912

RESULT 15

US-09-277-716-30

; Sequence 30, Application US/09277716A

; Patent No. 6232107

; GENERAL INFORMATION:

; APPLICANT: Bryan, Bruce

; APPLICANT: Szent-Gyorgyi, Christopher

; APPLICANT: PROLUME, LTD.

; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE

; CURRENT APPLICATION NUMBER: US/09/277,716A

; CURRENT FILING DATE: 1999-03-26

; EARLIER APPLICATION NUMBER: 60/102,939

; EARLIER FILING DATE: 1998-10-01

; EARLIER APPLICATION NUMBER: 60/089,367

; EARLIER FILING DATE: 1998-06-15

; EARLIER APPLICATION NUMBER: 60/079,624

; EARLIER FILING DATE: 1998-03-27

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 30

; LENGTH: 1104

; TYPE: DNA

; ORGANISM: *Ptilosarcus gurneyi*

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (34)..(747)

; FEATURE:

; OTHER INFORMATION: *Ptilosarcus* Green Fluorescent Protein (GFP) (insect A)

US-09-277-716-30

```

US-10-006-922A-12 (1-225) x US-09-277-716-30 (1-1104)

QY      2 ArgSerSerLysAsnVal-----IleLysGluPheMetArgPheLysVal 16
      :: :::::||||| ::::: ||||| ||||| |||||
Db      31 AAAATGAACCGCAACGTATTAAAGAACACTGGACTGAAAGAGATTATGTCGGCAAAAGCT 90

QY      17 ArgMetGluGlyThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArg 36
      :::::||||| ||||| ||||| ||||| :::::||||| ||||| |||||
Db      91 AGCGTTGAAGGAATCGTGAACAATCACGTTTTTTCATGGAAGGATTGTGAAAAAGGCAAT 150

QY      37 ProTyrGluGlyHisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPhe 56
      |||::: :::::||||| ||||| ||||| ||||| ||||| |||||
Db      151 GTATTATTTGGAAACCAATTGATGCAAAATCCGGGTTACAAAGGGAGGTCCGTTGCCATTTC 210

QY      57 AlaTrpAspIleLeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisPro 76
      |||:::|||||::: ||||| ||||| ||||| ::::: ||||| |||||
Db      211 GCTTTTCGATATTGTTTCCATAGCTTTTCCAATACGGGAATCGCACTTTTCACGAAATACCCA 270

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2005, 14:24:39 ; Search time 3345 Seconds  
(without alignments)  
9821.416 Million cell updates/sec

Title: US-10-006-922A-11  
Perfect score: 678  
Sequence: 1 atgagggtttccaagaatgt.....gccaccatctgttccttaa 678

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	678	100.0	678	AR183915	AR183915 Sequence
2	678	100.0	678	AX172854	AX172854 Sequence
3	678	100.0	678	AX207715	AX207715 Sequence
4	678	100.0	678	AX233581	AX233581 Sequence
5	678	100.0	859	AF168419	AF168419 Discosoma
6	678	100.0	859	AX463698	AX463698 Sequence
7	666.8	98.3	678	AF545828	AF545828 Discosoma
8	665	98.1	666	AX348043	AX348043 Sequence
9	665	98.1	666	AX353910	AX353910 Sequence
10	661	97.5	711	AY679107	AY679107 Discosoma
11	658.8	97.2	898	AX686888	AX686888 Sequence
12	656.2	96.8	921	AY679106	AY679106 Discosoma
13	566.4	83.5	876	AF272711	AF272711 Discosoma
14	566.4	83.5	876	AX686894	AX686894 Sequence
15	450.6	66.5	723	AR527331	AR527331 Sequence
16	449.8	66.3	691	AX233584	AX233584 Sequence
17	449.8	66.3	713	AX233627	AX233627 Sequence
18	449.2	66.3	678	AX370404	AX370404 Sequence
19	449.2	66.3	678	AX824725	AX824725 Sequence

20	447.2	66.0	2721	6	CQ882115	CQ882115 Sequence
21	447.2	66.0	2772	6	CQ882117	CQ882117 Sequence
22	447.2	66.0	4692	6	AX463702	AX463702 Sequence
23	447.2	66.0	9320	6	AX663075	AX663075 Sequence
24	446	65.8	678	6	AX370406	AX370406 Sequence
25	445.8	65.8	6893	6	AX823860	AX823860 Sequence
26	445.6	65.7	8811	12	AY569779	AY569779 Cloning v
27	445.6	65.7	12404	12	AY569780	AY569780 Cloning v
28	439.6	64.8	678	6	AX370408	AX370408 Sequence
29	438	64.6	678	6	AX824732	AX824732 Sequence
30	437	64.5	1050	6	AX666133	AX666133 Sequence
31	437	64.5	4488	6	CQ849509	CQ849509 Sequence
32	437	64.5	6423	12	AY613997	AY613997 Cloning v
33	437	64.5	7616	6	CQ849511	CQ849511 Sequence
34	435.6	64.2	675	6	AX824731	AX824731 Sequence
35	434.6	64.1	850	6	AX686886	AX686886 Sequence
36	432.2	63.7	10141	12	AY342347	AY342347 Red H-Pel
37	432.2	63.7	10276	12	AY342348	AY342348 Red H-Sti
38	432.2	63.7	10481	12	AY490568	AY490568 UAS-Red S
39	430	63.4	871	3	AY679108	AY679108 Discosoma
40	410.4	60.5	681	12	AF506025	AF506025 Synthetic
41	410.2	60.5	1395	12	AF506026	AF506026 Synthetic
42	409.6	60.4	699	6	AR183914	AR183914 Sequence
43	409.6	60.4	699	6	AX172853	AX172853 Sequence
44	409.6	60.4	699	6	AX207714	AX207714 Sequence
45	409.6	60.4	699	6	AX233580	AX233580 Sequence

ALIGNMENTS

RESULT 1  
AR183915  
LOCUS AR183915 678 bp DNA linear PAT 20-APR-2002  
DEFINITION Sequence 6 from patent US 6342379.  
ACCESSION AR183915  
VERSION AR183915.1 GI:20227884  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 678)  
AUTHORS Tsien,R.Y. and Gonzalez,J.B. III.  
TITLE Detection of transmembrane potentials by optical methods  
JOURNAL Patent: US 6342379-A 6-29-JAN-2002;  
FEATURES Location/Qualifiers  
source 1..678  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match	100.0%;	Score 678;	DB 6;	Length 678;
Best Local Similarity	100.0%;	Pred. No. 1.2e-189;		
Matches 678;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTTCATGAGGTTTAAAGTTTCGCATGGAAGGA	60	
Db	1	ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTTCATGAGGTTTAAAGTTTCGCATGGAAGGA	60	
Qy	61	ACGGTCAATGGGCACGAGTTTGAATAAGCGGAGAGGGAGGCCATACGAAGGC	120	
Db	61	ACGGTCAATGGGCACGAGTTTGAATAAGCGGAGAGGGAGGCCATACGAAGGC	120	
Qy	121	CACATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTGCTTGGGATATT	180	
Db	121	CACATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTGCTTGGGATATT	180	
Qy	181	TTGTACCAACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA	240	
Db	181	TTGTACCAACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA	240	
Qy	241	GACTATAAAAAGCTGTCAATTTCCCTGAAGGATTTAAATGGGAAGGGTTCATGAACCTTTGAA	300	

best primer and under 102e

Db	241	GACTATAAAAAGCTGTCAATTCCTGAAGGATTTAAATGGAAAGGGTCATGAACCTTTGAA	300
QY	301	GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTCAGGATGGCTGTTTCATCTAC	360
Db	301	GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTCAGGATGGCTGTTTCATCTAC	360
QY	361	AAGGTCAGGTTCAATGGCGTGAACCTTTCCTCCGATGGACCTGTTATGCAAAAAGAACA	420
Db	361	AAGGTCAGGTTCAATGGCGTGAACCTTTCCTCCGATGGACCTGTTATGCAAAAAGAACA	420
QY	421	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGATCCTCGTGATGGCGTGTGAAAGGAGAG	480
Db	421	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGATCCTCGTGATGGCGTGTGAAAGGAGAG	480
QY	481	ATTATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT	540
Db	481	ATTATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT	540
QY	541	TACATGGCAAAGAAGCCCTGTGCAGCTACCAGGTACTACTATGTTGACTCCAAACTGGAT	600
Db	541	TACATGGCAAAGAAGCCCTGTGCAGCTACCAGGTACTACTATGTTGACTCCAAACTGGAT	600
QY	601	ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCGAGGGACGC	660
Db	601	ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCGAGGGACGC	660
QY	661	CACCATCTGTTCCCTTTAA	678
Db	661	CACCATCTGTTCCCTTTAA	678

RESULT 2

AX172854

LOCUS AX172854 678 bp DNA linear PAT 03-JUL-2001

DEFINITION Sequence 6 from Patent WO0142211.

ACCESSION AX172854

VERSION AX172854.1 GI:14597903

KEYWORDS

SOURCE

ORGANISM

1

Discosoma sp.

Discosoma sp.

Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia; Discosomatidae; Discosoma.

REFERENCE

AUTHORS Tsien, R. Y. and Gonzalez, J. E.

TITLE Detection of transmembrane potentials by optical methods

JOURNAL Patent: WO 0142211-A 6 14-JUN-2001;

The Regents of the University of California (US)

LOCATION/Qualifiers

1..678

/organism="Discosoma sp."

/mol\_type="unassigned DNA"

/db\_xref="taxon:86600"

FEATURES

source

ORIGIN

Handwritten signature: *Handwritten signature*

QY	241	GACTATAAAAAGCTGTGTCATTTCTCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAA	300
Db	241		300
QY	301	GACGGTGGCGTCGTTACTGTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC	360
Db	301		360
QY	361	AAGGTCAAGTTCATTTGGCGTGAACCTTCCCTCCGATGGACCTGTATTGCAAAAAGAGACA	420
Db	361		420
QY	421	ATGGGCTGGGAAGCCAGCACACTGAGCGTTTGTATCCTCGTGATGGCGTGTGTAAAGGAGAG	480
Db	421		480
QY	481	ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT	540
Db	481		540
QY	541	TACATGGCAAAAGAACCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT	600
Db	541		600
QY	601	ATAACAAGCCACAAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC	660
Db	601		660
QY	661	CACCATCTGTTCCCTTTAA	678
Db	661		678

RESULT 3  
 AX207715  
 LOCUS AX207715 678 bp DNA linear  
 DEFINITION Sequence 13 from Patent WO0157242.  
 ACCESSION AX207715  
 VERSION AX207715.1 GI:15422399  
 KEYWORDS  
 SOURCE Discosoma sp.  
 ORGANISM Discosoma sp.  
 Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;  
 Corallimorpharia; Discosomatidae; Discosoma.  
 1  
 REFERENCE  
 AUTHORS Stack, J.H., Whitney, M., Cubitt, A.B. and Pollok, B.A.  
 TITLE Methods of protein destabilization and uses thereof  
 JOURNAL Patent: WO 0157242-A 13 09-AUG-2001;  
 Aurora Biosciences Corporation (US)  
 FEATURES  
 source  
 1. .678  
 /organism="Discosoma sp."  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:86600"

Accession	070,	Conservative	0,	Matches	0,	Indels	0,	Gaps
QY	1	ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTTCATGAGGTTTTAAGGTTTCGCATGGAAGGA	60					
Db	1	ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTTCATGAGGTTTTAAGGTTTCGCATGGAAGGA	60					
QY	61	ACGGTCAATGGGCACGAGTTTGAATATGAAGGCGAAGGAGGGGGAGGCCATACGAAAGGC	120					
Db	61	ACGGTCAATGGGCACGAGTTTGAATATGAAGGCGAAGGAGGGGGAGGCCATACGAAAGGC	120					
QY	121	CACAAATACCGTAAAGCTTAAGGTAACCAAGGGGGGACCTTTGCCATTTTGCTTGGGATATT	180					
Db	121	CACAAATACCGTAAAGCTTAAGGTAACCAAGGGGGGACCTTTGCCATTTTGCTTGGGATATT	180					
QY	181	TTGTCACCAACAATTTTCAGTATGGAAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA	240					
Db	181	TTGTCACCAACAATTTTCAGTATGGAAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA	240					





FEATURES	Location/Qualifiers	
	source	1. .859
		/organism="Discosoma sp."
		/mol_type="mRNA"
		/db_xref="taxon:86600"
CDS		54. .731
		/note="GFP-like protein; orange-red"
		/codon_start=1
		/product="fluorescent protein FP583"
		/protein_id="AAF03369.1"
ORIGIN		/db_xref="GI:6090867"
		/translation="MRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGERPYEGHNTV
		KLKVTGGPLPFAWDILSPQFYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDG
		GVVTVDSSLDGCFIYKVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGE
		IHKALKLDGGHYLVEFKSIYMAKPVQLPGYVVVDSKLDITSHNEDYTIVEQYERTE
GRHHLFL"		
Query Match 100.0%; Score 678; DB 3; Length 859;		
Best Local Similarity 100.0%; Pred. No. 1.2e-189;		
Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTGCGATGGAAGGA 60
Db	54	ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTGCGATGGAAGGA 113
QY	61	ACGGTCAATGGGCACGAGTTTGAATAAGAGCGAAGGAGAGGGGAGGCATACGAAGGC 120
Db	114	ACGGTCAATGGGCACGAGTTTGAATAAGAGCGAAGGAGAGGGGAGGCATACGAAGGC 173
QY	121	CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTCCTTGGGATATT 180
Db	174	CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTCCTTGGGATATT 233
QY	181	TTGTCACCACAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240
Db	234	TTGTCACCACAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 293
QY	241	GACTATAAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300
Db	294	GACTATAAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 353
QY	301	GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360
Db	354	GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 413
QY	361	AAGGTCAAGTTCAATGGCGTGAACCTTCCGATGGACCTGTTGATGAAATCAAAGATATT 480
Db	474	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTATGGCGTGTGAAAGGAGAG 533
QY	481	ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATCAAAGATATT 540
Db	534	ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATCAAAGATATT 593
QY	541	TACATGGCAAGAAGCCTGTGCAGCTACCAGGCTACTACTATGTTGACTCCAAACTGGAT 600
Db	594	TACATGGCAAGAAGCCTGTGCAGCTACCAGGCTACTACTATGTTGACTCCAAACTGGAT 653
QY	601	ATAACAAGCCACACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGGAGGC 660
Db	654	ATAACAAGCCACACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGGAGGC 713
QY	661	CACCATCTGTTCCCTTTAA 678
Db	714	CACCATCTGTTCCCTTTAA 731
RESULT 6		
AX463698		

LOCUS	AX463698	Sequence 12 from Patent WO0248338.	859 bp	DNA	linear	PAT 15-JUL-2000
DEFINITION	AX463698					
ACCESSION	AX463698					
VERSION	AX463698.1	GI:21886457				
KEYWORDS						
SOURCE	Discosoma sp.					
ORGANISM	Discosoma sp.					
REFERENCE						
AUTHORS	Lichtenberg-Frat,H.					
TITLE	Yeast strain for testing the geno- and cytotoxicity of complex environmental contamination					
JOURNAL	Patent: WO 0248338-A 12 20-JUN-2002;					
	Lichtenberg-Frat, Hella (DE)					
FEATURES	Location/Qualifiers					
source	1. .859					
	/organism="Discosoma sp."					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:86600"					
CDS	54. .731					
	/note="unnamed protein product"					
	/codon_start=1					
	/protein_id="CAD42147.1"					
	/db_xref="GI:21886458"					
	/translation="MRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGERPYEGHNTV KLKVTGGPLPFAWDILSPQFYGSKVYVKHPADI PDYKKLSFPEGFKWERVMNFEDG GVVTVDSSLDGCFIYKVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGE IHKALKLDGGHYLVEFKSIYMAKPVQLPGYVVVDSKLDITSHNEDYTIVEQYERTE GRHHLFL"					
ORIGIN						
Query Match	100.0%;	Score 678;	DB 6;	Length 859;		
Best Local Similarity	100.0%;	Pred. No. 1.2e-189;				
Matches	678;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	1	ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTTCGCATGGAAGGA	60			
Db	54	ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTTCGCATGGAAGGA	113			
Qy	61	ACGGTCAATGGGCACGAGTTTGAATAAGAGCGAAGGAGAGGGGAGGCATACGAAGGC	120			
Db	114	ACGGTCAATGGGCACGAGTTTGAATAAGAGCGAAGGAGAGGGGAGGCATACGAAGGC	173			
Qy	121	CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT	180			
Db	174	CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT	233			
Qy	181	TTGTCACCACAAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA	240			
Db	234	TTGTCACCACAAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA	293			
Qy	241	GACTATAAAAGCTGTCAATTTCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA	300			
Db	294	GACTATAAAAGCTGTCAATTTCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA	353			
Qy	301	GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC	360			
Db	354	GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC	413			
Qy	361	AAGGTCAAGTTCAATGGCGTGAACCTTTCTCCGATGGACCTGTTATGCAAAAAGAAGACA	420			
Db	414	AAGGTCAAGTTCAATGGCGTGAACCTTTCTCCGATGGACCTGTTATGCAAAAAGAAGACA	473			
Qy	421	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTATGGCGTGTGAAAGGAGAG	480			
Db	474	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTATGGCGTGTGAAAGGAGAG	533			
Qy	481	ATTCAATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT	540			
Db	534	ATTCAATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT	593			
Qy	541	TACATGGCAAGAAGCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT	600			

```
|||||
Db 594 TACATGGCAAGAAGCCTGTGCAGCTACAGGGTACTACTATGTTGACTCCAAACTGGAT 653
|||||
QY 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCGAGGGACGC 660
|||||
Db 654 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCGAGGGACGC 713
|||||
QY 661 CACCATCTGTTCCCTTTAA 678
|||||
Db 714 CACCATCTGTTCCCTTTAA 731

RESULT 7
AF545828
LOCUS AF545828 678 bp mRNA linear INV 04-JUN-2004
DEFINITION Discosoma sp. JW-2002 orange fluorescent protein FP586 mRNA,
complete cds.
ACCESSION AF545828
VERSION AF545828.1 GI:33333763
KEYWORDS Discosoma sp. JW-2002
SOURCE Discosoma sp. JW-2002
ORGANISM Discosoma sp. JW-2002
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;
Corallimorpharia; Discosomatidae; Discosoma.
REFERENCE 1 (bases 1 to 678)
AUTHORS Wiedemann,J. and Girod,A.
TITLE Direct Submission
JOURNAL Submitted (12-SEP-2002) Abteilung Allgemeine Zoologie und
Endokrinologie, Universitaet Ulm, Albert Einstein Allee 11, Ulm
89069, Germany
FEATURES
source
Location/Qualifiers
1..678
/organism="Discosoma sp. JW-2002"
/mol_type="mRNA"
/db_xref="taxon:208461"
/note="isolated from specimen with smooth disc"
1..678
/note="GFP-like protein, orange-red; dsFP586; similar to
fluorescent protein dsFP583 (dsRed)"
/codon_start=1
/product="orange fluorescent protein FP586"
/protein_id="AAQ11987.1"
/db_xref="GI:33333764"
/translation="MSCSKXNVIKFPMPKVRMEGTVNGHEFEIEGEGEGRPYGGHNTV
KLKVTKGGLPFPFADWILSPQFYQSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDG
GVVTVQDSSLQDGCFIYKVFIGNFSPDGPVMQKKTMGWEASTERLYPRDGLKGD
IHKALKLKDGGHYLVFEKSIYMAKKPVQLPGYIYVDSKLDITSHNEDYTIQEYVERE
GRHHLFL"
ORIGIN
Query Match 98.3%; Score 666.8; DB 3; Length 678;
Best Local Similarity 99.0%; Pred. No. 2.5e-186;
Matches 671; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 ATGAGGCTTCCAAGATGTTATCAAGGAGTTCATGAGGTTTAAGGTTCCGATGGAAGGA 60
|||||
Db 1 ATGAGTTGTTCCAAGATGTTATCAAGGAGTTCATGAGGTTTAAGGTTCCGATGGAAGGA 60
|||||
QY 61 ACGGTCAATGGGACGAGTTTGAATAGAGCGAAGGAGAGGGGAGGCATACGAAGGC 120
|||||
Db 61 ACGGTCAATGGGACGAGTTTGAATAGAGCGAAGGAGAGGGGAGGCATACGGAGGC 120
|||||
QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTGGGATATT 180
|||||
Db 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTGGGATATT 180
|||||
QY 181 TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240
|||||
Db 181 TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240
|||||
QY 241 GACTATAAAAAGCTGTCTATTTCTCTGAAGGATTTAAATGGGAAAGGGTCACTGAACCTTTGAA 300
|||||
Db 241 GACTATAAAAAGCTGTCTATTTCTCTGAAGGATTTAAATGGGAAAGGGTCACTGAACCTTTGAA 300
|||||
```

```
QY 301 GACGGTGGCGTCGTTACTGTAAACCAGGATTCCAGTTTGCAGGATGGCTGTTTCATCTAC 360
|||||
Db 301 GACGGTGGCGTCGTTACTGTAAACCAGGATTCCAGTTTGCAGGATGGCTGTTTCATCTAC 360
|||||
QY 361 AAGGTCAAGTTTCAATTGGCGTGAACTTTCCTTCCGATGGACCTGTATTGCAAAAAGAAGACA 420
|||||
Db 361 AAGGTCAAGTTTCAATTGGCGTGAACTTTCCTTCCGATGGACCTGTATTGCAAAAAGAAGACA 420
|||||
QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGATCCTCGTGATGGCGTGTGAAAAGGAGAG 480
|||||
Db 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGATCCTCGTGATGGCGTGTGAAAAGGAGAG 480
|||||
QY 481 ATTATAAGGCTCTGAAGCTGAAAGACGGTGGTCAATTACCTAGTTGAATTCAAAAAGTATT 540
|||||
Db 481 ATTATAAGGCTCTGAAGCTGAAAGACGGTGGTCAATTACCTAGTTGAATTCAAAAAGTATT 540
|||||
QY 541 TACATGGCAAAAGACCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600
|||||
Db 541 TACATGGCAAAAGACCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600
|||||
QY 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCGAGGGACGC 660
|||||
Db 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCGAGGGACGC 660
|||||
QY 661 CACCATCTGTTCCCTTTAA 678
|||||
Db 661 CACCATCTGTTCCCTTTAA 678
|||||

RESULT 8
AX348043
LOCUS AX348043 666 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 22 from Patent EP1172375.
ACCESSION AX348043
VERSION AX348043.1 GI:18614153
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Odenthal,M. and Jung,D.
TITLE Gene expression, genome alteration and reporter gene expression in
myofibroblasts and myofibroblasts-like cells by using the
regulatory regions within the alpha smooth muscle alpha-actin gene
Patent: EP 1172375-A 22 16-JAN-2002;
JOURNAL Odenthal, Margarete (DE)
FEATURES
Location/Qualifiers
1..666
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Red Fluorescent Protein"
ORIGIN
Query Match 98.1%; Score 665; DB 6; Length 666;
Best Local Similarity 100.0%; Pred. No. 8.4e-186;
Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 13 AAGAATGTTATCAAGGAGTTTCATGAGGTTTAAAGGTTTCGCATGGAAGGACGGTCAATGGG 72
|||||
Db 1 AAGAATGTTATCAAGGAGTTTCATGAGGTTTAAAGGTTTCGCATGGAAGGACGGTCAATGGG 60
|||||
QY 73 CACGAGTTTGAAATAGAAGCGGAAGGAGAGGGGAGGCCATACGAAGGCCACAATACCGTA 132
|||||
Db 61 CACGAGTTTGAAATAGAAGCGGAAGGAGAGGGGAGGCCATACGAAGGCCACAATACCGTA 120
|||||
QY 133 AAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTGGGATATTTGTCAACCAAA 192
|||||
Db 121 AAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTGGGATATTTGTCAACCAAA 180
|||||
QY 193 TTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCAACTATAAAAAG 252
|||||
```

Db 181 TTTCAGTATGGAACGAAGGTATATGTCAAGCACCCCTGCCGACATACCAGACTATAAAAAAG 240

QY 253 CTGTCAATTTCCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTTGAAGACGGTGGCGTC 312

Db 241 CTGTCAATTTCCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTTGAAGACGGTGGCGTC 300

QY 313 GTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTACAAGGTCAAGTTC 372

Db 301 GTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTACAAGGTCAAGTTC 360

QY 373 ATTGGCGTGAACTTTCCTCCGATGGACCTGTATGCAAAAGAACACAAATGGGCTGGGAA 432

Db 361 ATTGGCGTGAACTTTCCTCCGATGGACCTGTATGCAAAAGAACACAAATGGGCTGGGAA 420

QY 433 GCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAGATTTCATAAGGCT 492

Db 421 GCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAGATTTCATAAGGCT 480

QY 493 CTGAAGCTGAAAGACCGTGGTCATTACCTAGTTGAATTCAAAAGTATTTCATACATGGCAAG 552

Db 481 CTGAAGCTGAAAGACCGTGGTCATTACCTAGTTGAATTCAAAAGTATTTCATACATGGCAAG 540

QY 553 AAGCCTGTGCAGTACCAGGGTACTACTATGTTGACTCCAAACTGGATATAACAAGCCAC 612

Db 541 AAGCCTGTGCAGTACCAGGGTACTACTATGTTGACTCCAAACTGGATATAACAAGCCAC 600

QY 613 AACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGCCACCATCTGTTC 672

Db 601 AACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGCCACCATCTGTTC 660

QY 673 CTTTA 677

Db 661 CTTTA 665

RESULT 9

AX353910

LOCUS AX353910

DEFINITION Sequence 22 from Patent WO0204509.

ACCESSION AX353910

VERSION AX353910.1 GI:18618888

KEYWORDS synthetic construct

SOURCE synthetic construct

ORGANISM other sequences; artificial sequences.

REFERENCE 1

AUTHORS Odenthal,M. and Jung,D.

TITLE Gene expression, genome alteration and reporter expression in myofibroblasts and myofibroblast-like cells

JOURNAL Patent: WO 0204509-A 22 17-JAN-2002;

Odenthal, Margarete (DE)

FEATURES

Location/Qualifiers

1..666

/organism="synthetic construct"

/mol\_type="unassigned DNA"

/db\_xref="taxon:32630"

/note="Red Fluorescent Protein"

ORIGIN

Query Match 98.1%; Score 665; DB 6; Length 666;

Best Local Similarity 100.0%; Pred. No. 8.4e-186;

Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 AAGAATGTTATCAAGGAGTTCATGAGGTTTAAAGTTTCGCATGGAAGGAACGGTCAATGGG 72

Db 1 AAGAATGTTATCAAGGAGTTCATGAGGTTTAAAGTTTCGCATGGAAGGAACGGTCAATGGG 60

QY 73 CACGAGTTTGAATAGAACGGCGAAGGAGAGGGAGGCCCATACGAAGGCCACAAATACCGTA 132

Db 61 CACGAGTTTGAATAGAACGGCGAAGGAGAGGGAGGCCCATACGAAGGCCACAAATACCGTA 120

QY 133 AAGCTTAAGGTAAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATTTTGTCAACCACAA 192

Db 121 AAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTTGCTTTGGGATATTTTGTCAACCACAA 180

QY 193 TTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCAGACTATAAAAAAG 252

Db 181 TTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCAGACTATAAAAAAG 240

QY 253 CTGTCAATTTCCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTTGAAGACGGTGGCGTC 312

Db 241 CTGTCAATTTCCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTTGAAGACGGTGGCGTC 300

QY 313 GTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTACAAGGTCAAGTTC 372

Db 301 GTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTACAAGGTCAAGTTC 360

QY 373 ATTGGCGTGAACTTTCCTCCGATGGACCTGTTATGCAAAAGAACACAAATGGGCTGGGAA 432

Db 361 ATTGGCGTGAACTTTCCTCCGATGGACCTGTTATGCAAAAGAACACAAATGGGCTGGGAA 420

QY 433 GCCAGCACTGAGCGTTCCTCGTGATGGCGTGTGAAAGGAGAGATTTCATAAGGCT 492

Db 421 GCCAGCACTGAGCGTTCCTCGTGATGGCGTGTGAAAGGAGAGATTTCATAAGGCT 480

QY 493 CTGAAGCTGAAAGACCGTGGTCATACCTAGTTGAATTCAAAAGTATTTCATACATGGCAAG 552

Db 481 CTGAAGCTGAAAGACCGTGGTCATACCTAGTTGAATTCAAAAGTATTTCATACATGGCAAG 540

QY 553 AAGCCTGTGCAGTACCAGGGTACTACTATGTTGACTCCAAACTGGATATAACAAGCCAC 612

Db 541 AAGCCTGTGCAGTACCAGGGTACTACTATGTTGACTCCAAACTGGATATAACAAGCCAC 600

QY 613 AACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGCCACCATCTGTTC 672

Db 601 AACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGCCACCATCTGTTC 660

QY 673 CTTTA 677

Db 661 CTTTA 665

RESULT 10

AY679107

LOCUS AY679107

DEFINITION Discosoma sp. RC-2004 enhanced red fluorescent protein R+ mRNA, complete cds.

ACCESSION AY679107

VERSION AY679107.1 GI:51472046

KEYWORDS

SOURCE Discosoma sp. RC-2004

ORGANISM Discosoma sp. RC-2004

REFERENCE 1 (bases 1 to 711)

AUTHORS Carter,R.W., Gibbs,P.D.L. and Schmale,M.C.

TITLE Cloning of Anthozoan Fluorescent Protein Genes

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 711)

AUTHORS Carter,R.W., Gibbs,P.D.L. and Schmale,M.C.

TITLE Direct Submission

JOURNAL Submitted (06-JUL-2004) MBF, University of Miami, 4600 Rickenbacker Cswy, Miami, FL 33129, USA

FEATURES

Location/Qualifiers

1..711

/organism="Discosoma sp. RC-2004"

/mol\_type="mRNA"

/db\_xref="taxon:289055"

1..711

/note="GFP-like; DspR+"

/codon\_start=1

/product="enhanced red fluorescent protein R+"

/protein\_id="AAU04444.1"

/db\_xref="GI:51472047"

CDS

/translation="MSCSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGERPYEGHNTV

KLKVTKGGLPFPAWDILSPQFYQGSKVYKHPADIPDYKLSFPEGFKWERVMNPFDG



GVVTVTQDSSLQDGCFIYKVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGVLKGE  
IHKALKLKDGGHYLVFEKTIYMAKFPVQLPGYIYVDSKLDITSHNKDYTIVEQYERTE  
GRHFLPLKAEELGSNVGER"

ORIGIN

Query Match 97.5%; Score 661; DB 3; Length 711;  
Best Local Similarity 98.5%; Pred. No. 1.3e-184;  
Matches 667; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy	1	ATGAGGTCCTTCCAAGAATGTTTATCAAGGAGTTCATGAGGTTTAAGGTTTCGCATGGAAGGA	60
Db	1	ATGAGTTGTTCCAAGAATGTTTATCAAGGAGTTCATGAGGTTTAAGGTTTCGCATGGAAGGA	60
Qy	61	ACGGTCAATGGGCACGAGTTTGAATAGAACGCGAAGGAGAGGGGAGGCCATACGAAGGC	120
Db	61	ACGGTCAATGGGCACGAGTTTGAATAGAACGCGAAGGAGAGGGGAGGCCATACGAAGGC	120
Qy	121	CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTCCTTGGGATATT	180
Db	121	CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTCCTTGGGATATT	180
Qy	181	TTGTCAACCAACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCTTCGCCGACATACCA	240
Db	181	TTGTCAACCAACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCATCTTCGCCGACATACCA	240
Qy	241	GACTATAAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA	300
Db	241	GACTATAAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA	300
Qy	301	GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC	360
Db	301	GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC	360
Qy	361	AAGGTCAAGTTTCATTTGGCGTGAACCTTTCCTCCGATGGACCTGTTATGCAAAAGAGACA	420
Db	361	AAGGTCAAGTTTCATTTGGCGTGAACCTTTCCTTCGATGGACCTGTTATGCAAAAGAGACA	420
Qy	421	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG	480
Db	421	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG	480
Qy	481	ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT	540
Db	481	ATTCATAAGGCTCTGAAGTTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT	540
Qy	541	TACATGGCAAGAAGCCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAT	600
Db	541	TACATGGCAAGAAGCCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAT	600
Qy	601	ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGGAGCGC	660
Db	601	ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGGAGCGC	660
Qy	661	CACCATCTGTTTCCTTTA	677
Db	661	CACCATCTGTTTCCTTAA	677

RESULT 11  
AX686888  
LOCUS AX686888 898 bp DNA linear PAT 31-MAR-2003  
DEFINITION Sequence 11 from Patent WO0127150.  
ACCESSION AX686888  
VERSION AX686888.1 GI:29409468  
KEYWORDS  
SOURCE Discosoma sp.  
ORGANISM Discosoma sp.  
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;  
Corallimorpharia; Discosomatidae; Discosoma.

REFERENCE 1  
AUTHORS Lukyanov, S.A., Fradkov, A.F., Labas, Y.A., Matz, M.V. and Terskikh, A.  
TITLE Anthozoa derived chromo/fluoroproteins and methods for using the same

JOURNAL Patent: WO 0127150-A 11 19-APR-2001;  
Clontech Laboratories Inc. (US)  
FEATURES Location/Qualifiers  
source 1..898  
/organism="Discosoma sp."  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:85600"

ORIGIN

Query Match 97.2%; Score 658.8; DB 6; Length 898;  
Best Local Similarity 98.2%; Pred. No. 5.9e-184;  
Matches 666; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy	1	ATGAGGTCCTTCCAAGAATGTTTATCAAGGAGTTCATGAGGTTTAAAGTTTCGCATGGAAGGA	60
Db	93	ATGAGGTCCTTCCAAGAATGTTTATCAAGGAGTTCATGAGGTTTAAAGTTTCGCATGGAAGGA	152
Qy	61	ACGGTCAATGGGCACGAGTTTGAATAGAACGCGAAGGAGAGGGGAGGCCATACGAAGGC	120
Db	153	ACGGTCAATGGGCACGAGTTTGAATAGAACGCGAAGGAGAGGGGAGGCCATACGAAGGC	212
Qy	121	CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTCCTTGGGATATT	180
Db	213	CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTCCTTGGGATATT	272
Qy	181	TTGTCAACCAACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCTTCGCCGACATACCA	240
Db	273	TTGTCAACCAACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCTTCGCCGACATACCA	332
Qy	241	GACTATAAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA	300
Db	333	GACTATAAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA	392
Qy	301	GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC	360
Db	393	GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC	452
Qy	361	AAGGTCAAGTTTCATTTGGCGTGAACCTTTCCTCCGATGGACCTGTTATGCAAAAGAGACA	420
Db	453	AAGTCAAGTTTCATTTGGCGTTGAACCTTTCCTCCGATGGACCTGTTATGCAAAAGAGACA	512
Qy	421	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG	480
Db	513	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG	572
Qy	481	ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT	540
Db	573	ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT	632
Qy	541	TACATGGCAAGAAGCCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAT	600
Db	633	TACATGGCAAGAAGCCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAT	692
Qy	601	ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGGAGCGC	660
Db	693	ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGGAGCGC	752
Qy	661	CACCATCTGTTTCCTTTAA	678
Db	753	CACCATCTGTTTCCTTTAA	770

RESULT 12  
AY679106  
LOCUS AY679106 921 bp mRNA linear INV 25-AUG-2004  
DEFINITION Discosoma sp. RC-2004 red fluorescent protein R1 mRNA, complete cds.  
ACCESSION AY679106  
VERSION AY679106.1 GI:51472044  
KEYWORDS  
SOURCE Discosoma sp. RC-2004  
ORGANISM Discosoma sp. RC-2004  
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;





QY 181 TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240  
Db 225 TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAACACCCCTGCCGACATACCA 284  
QY 241 GACTATAAAAGCTGTCAATTTCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAA 300  
Db 285 GACTATAAAAGCTGTCAATTTCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAA 344  
QY 301 GACGGTGGCGTCTTACTGTAAACCCAGGATTCAGTTTGACAGGATGGCTGTTTCATCTAC 360  
Db 345 GACGGTGGCGTCTTACTGTATCCCAAGATTCAGTTTGAAAGACGGCTGTTTCATCTAC 404  
QY 361 AAGGTCAAGTTCAATGGCGTGAACTTTCCCTCCGATGGACCTGTTATGCAAAAGAGACA 420  
Db 405 GAGGTCAAGTTCAATGGCGTGAACTTTCCCTCCGATGGACCTGTTATGCAAGGAGACA 464  
QY 421 ATGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
Db 465 CGGGCTGGGAAGCCAGCTCTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 524  
QY 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGCTCAATACCTAGTTGAATTCAAAAGTATT 540  
Db 525 ATCCATATGGCTCTGAGGCTGGAAGGAGCGGCCATTACCTCGTTGAATTCAAAAGTATT 584  
QY 541 TACATGGCAAGAGAGCCT---GTGCAGCTACAGGGTACTACTATGTTGACTCCAAACTG 597  
Db 585 TACATGGTAAAGAGCCTTCAGTGCAGTTGCCAGGCTACTATTATGTTGACTCCAAACTG 644  
QY 598 GATATAACAGCCACAACGAAGACTATACAAATCGTTGAGCAGTATGAAAGAACCGAGGGA 657  
Db 645 GATATGACGAGCCACAACGAAGATTACACAGTCGTTGAGCAGTATGAAAGAACCCAGGGA 704  
QY 658 CGCCACCATCTGTTCTCTTA 677  
Db 705 CGCCACCATCCGTTTCATTAA 724

RESULT 14  
AX686894  
LOCUS AX686894 876 bp DNA linear PAT 31-MAR-2003  
DEFINITION Sequence 17 from Patent WO0127150.  
ACCESSION AX686894  
VERSION AX686894.1 GI:29409472  
KEYWORDS Discosoma sp.  
SOURCE Discosoma sp.  
ORGANISM Discosoma sp.  
REFERENCE 1 Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia; Discosomatidae; Discosoma.  
AUTHORS Lukyanov,S.A., Pradkov,A.F., Labas,Y.A., Matz,M.V. and Terskikh,A.  
TITLE Anthozoa derived chromo/fluoroproteins and methods for using the same  
JOURNAL Patent: WO 0127150-A 17 19-APR-2001;  
Clontech Laboratories Inc. (US)  
FEATURES Location/Qualifiers  
source 1. 876  
/organism="Discosoma sp."  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:86600"

ORIGIN

Query Match 83.5%; Score 566.4; DB 6; Length 876;  
Best Local Similarity 90.6%; Pred. No. 1.5e-156;  
Matches 616; Conservative 0; Mismatches 61; Indels 3; Gaps 1;  
QY 1 ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTTCGTCATGGAAGGA 60  
Db 45 ATGAGTTGTTCCAAGAATGTGATCAAGGAGTTCATGAGGTTCAAGGTTTCGTCATGGAAGGA 104  
QY 61 ACGGTCAATGGGCACGAGTTTGAATAGAGCGAAGGAGAGGGGAGGCCATACGAAGGC 120  
Db 105 ACGGTCAATGGGCACGAGTTTGAATAGAGCGAAGGTGAAGGGAGGSCCTTACGAAGGT 164

QY 121 CACAATACCGTAAGCTTAAGGTAACCAAGGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
Db 165 CACTGTTCCGTAAGCTTATGGTAACCAAGGGTGGACCTTTGCCATTTGCTTGGGATATT 224  
QY 181 TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240  
Db 225 TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAACACCCCTGCCGACATACCA 284  
QY 241 GACTATAAAAGCTGTCAATTTCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAA 300  
Db 285 GACTATAAAAGCTGTCAATTTCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAA 344  
QY 301 GACGGTGGCGTCTTACTGTAAACCCAGGATTCAGTTTGACAGGATGGCTGTTTCATCTAC 360  
Db 345 GACGGTGGCGTCTTACTGTATCCCAAGATTCAGTTTGAAAGACGGCTGTTTCATCTAC 404  
QY 361 AAGGTCAAGTTCAATGGCGTGAACTTTCCCTCCGATGGACCTGTTATGCAAAAGAGACA 420  
Db 405 GAGGTCAAGTTCAATGGGGTGAACTTTCTCTGATGGACCTGTTATGCAAGGAGACA 464  
QY 421 ATGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
Db 465 CGGGCTGGGAAGCCAGCTCTGAGCGTTTGTATCCTCGTGATGGGGTGTGAAAGGAGAG 524  
QY 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGCTCAATACCTAGTTGAATTCAAAAGTATT 540  
Db 525 ATCCATATGGCTCTGAGGCTGGAAGGAGCGGCCATTACCTCGTTGAATTCAAAAGTATT 584  
QY 541 TACATGGCAAGAGCCT---GTGCAGCTACAGGGTACTACTATGTTGACTCCAAACTG 597  
Db 585 TACATGGTAAAGAGCCTTCAGTGCAGTTGCCAGGCTACTATTATGTTGACTCCAAACTG 644  
QY 598 GATATAACAGCCACAACGAAGACTATACAAATCGTTGAGCAGTATGAAAGAACCGAGGGA 657  
Db 645 GATATGACGAGCCACAACGAAGATTACACAGTCGTTGAGCAGTATGAAAGAACCCAGGGA 704  
QY 658 CGCCACCATCTGTTCTCTTA 677  
Db 705 CGCCACCATCCGTTTCATTAA 724

RESULT 15

AR527331  
LOCUS AR527331 723 bp DNA linear PAT 08-OCT-2004  
DEFINITION Sequence 1 from patent US 6723537.  
ACCESSION AR527331  
VERSION AR527331.1 GI:53914309  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 723)  
AUTHORS Peelle,B.  
TITLE Directed evolution of protein in mammalian cells  
JOURNAL Patent: US 6723537-A 1 20-APR-2004;  
FEATURES Location/Qualifiers  
source 1. 723  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Query Match 66.5%; Score 450.6; DB 6; Length 723;  
Best Local Similarity 79.3%; Pred. No. 3.3e-122;  
Matches 534; Conservative 0; Mismatches 139; Indels 0; Gaps 0;  
QY 2 TGAGGTCTTCCAAGAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTTCGTCATGGAAGGA 61  
Db 5 TGCCTCTTCCAAGAACGTCATCAAGGAGTTCATGCGCTTCAAGGTGCGCATGGAGGGA 64  
QY 62 CCGTCAATGGGCACGAGTTTGAATAGAGCGAAGGAGAGGGGAGGCCATACGAAGGCC 121  
Db 65 CCGTGAACGCCACGAGTTTCGAGATCGAGGGCGAGGGCGGCGCCCTTACGAGGGCC 124

Qy	122	ACAATACCGTAAAGCTTAAGGTAACCAAGSGGGACCTTTGCCATTTGCTTGGGATATTT	181
Db	125	ACAACACCGTGAAGCTGAAGGTGACCAAGSGGGCCCCCTGCCCTTCGCCTGGGACATCC	184
Qy	182	TGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCAG	241
Db	185	TGTCCCCCAGTTCAGTACGGCTCCAAGGTGTACGTGAAGCACCCCGCCGACATCCCCG	244
Qy	242	ACTATAAAAGCTGTCTATTTCCCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAAG	301
Db	245	ACTACAAGAAGCTGTCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGAACCTTCGAGG	304
Qy	302	ACGGTGGCGTCTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTTTTCATCTACA	361
Db	305	ACGGCGGCGTGTGACCGTGACCCAGGACTCTCCCTGCAGGACGGCTGCTTCATCTACA	364
Qy	362	AGGTCAAGTTTCATTGGCGTGAACCTTTCCTTCCGATGGACCTGTTATGCAAAAAGACAA	421
Db	365	AGGTGAAGTTTCATCGGCGTGAACCTTCCCTCCGACGGCCCCGTATGCAAGAAGACCA	424
Qy	422	TGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTATGGCGTGTGAAAGGAGAGA	481
Db	425	TGGGCTGGGAGGCCCTCCACCGAGCGCTGTACCCCCGCGACGGCGTGTGAAAGGCGGAGA	484
Qy	482	TTCATAGGCTCTGAAGCTGAAAGACGGTGGTCAATTACCTAGTTGAATTCAAAGATATTT	541
Db	485	TCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACCTGGTGGAGTTCAGAGATATCT	544
Qy	542	ACATGGCAAGAAGCCTGTGCAGCTACCAAGGTACTACTAIGTTGACTCCAAACTGGATA	601
Db	545	ACATGGCCAAGAAGCCCGTGCAGCTGCCCGGTACTACTACGTGGACTCCAAGCTGGACA	604
Qy	602	TAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCAGGGACGCC	661
Db	605	TCACCTCCCAACAACGAGGACTACACCATCGTGGAGCAGTACGAGCGCACCCGAGGCCGCC	664
Qy	662	ACCATCTGTTCTT	674
Db	665	ACCACCTGTTCTT	677

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2005, 14:18:00 ; Search time 3040 Seconds  
(without alignments)  
8489.334 Million cell updates/sec

Title: US-10-006-922A-11  
Perfect score: 678  
Sequence: 1 atgagggtctccaagaatgt.....gccaccatctgttcctttaa 678

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST: \*  
1: gb\_est1: \*  
2: gb\_est2: \*  
3: gb\_hlc: \*  
4: gb\_est3: \*  
5: gb\_est4: \*  
6: gb\_est5: \*  
7: gb\_est6: \*  
8: gb\_gss1: \*  
9: gb\_gss2: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score		Query Match		Length	DB ID	Description
1	50	7.4	671	4	BI381780	BFLG1_003	BI381780 BFLG1_003
2	49	7.2	681	4	BI386888	BFL26_001	BI386888 BFL26_001
3	48.4	7.1	561	4	BI379061	BFLG1_000	BI379061 BFLG1_000
4	48.4	7.1	577	4	BI386955	BFL26_001	BI386955 BFL26_001
5	45.8	6.8	459	4	BI386952	BFL26_001	BI386952 BFL26_001
6	45.2	6.7	591	4	BI382373	BFLG2_000	BI382373 BFLG2_000
7	44.8	6.6	707	4	BI382442	BFLG2_000	BI382442 BFLG2_000
8	43.6	6.4	537	4	BI376909	BFLG3_000	BI376909 BFLG3_000
9	43.2	6.4	629	4	BI384961	BFLG2_003	BI384961 BFLG2_003
10	41.2	6.1	930	9	CL065858	CH216-105	CL065858 CH216-105
11	39.8	5.9	472	2	BB854015	BB854015	BB854015 BB854015
12	39.2	5.8	554	4	BI815946	PfESToab3	BI815946 PfESToab3
13	39	5.8	592	1	AV719715	AV719715	AV719715 AV719715
14	39	5.8	808	8	BZ131346	CH230-345	BZ131346 CH230-345
15	38.6	5.7	571	5	BQ451564	PfESToab0	BQ451564 PfESToab0
16	38.6	5.7	580	5	BU496978	PfESToab6	BU496978 PfESToab6
17	38.6	5.7	600	5	BU497757	PfESToab8	BU497757 PfESToab8
18	38.6	5.7	604	5	BQ597249	PfESToab3	BQ597249 PfESToab3
19	38.6	5.7	607	5	BQ739542	PfESToab4	BQ739542 PfESToab4
20	38.6	5.7	616	5	BQ451730	PfESToab0	BQ451730 PfESToab0
21	38.4	5.7	571	4	BI387406	BFL26_002	BI387406 BFL26_002
22	38.4	5.7	687	4	BI387917	BFL26_002	BI387917 BFL26_002
23	38.2	5.6	629	4	BI380008	BFLG1_001	BI380008 BFLG1_001
24	38	5.6	480	9	CE292671	tigr-gss-	CE292671 tigr-gss-

25	37.8	5.6	797	9	CNS003F8	AL064634 Drosophil
c 26	37.8	5.6	1025	9	CNS07CE9	AL439063 T3 end of
27	37.6	5.5	1168	8	CC208512	CC208512 CH261-111
28	37.2	5.5	548	4	BM521209	BM521209 sal35h10.
c 29	37.2	5.5	558	6	CD414787	CD414787 Gm ck4789
30	37.2	5.5	704	4	BG788189	BG788189 SEUMC008
31	37.2	5.5	888	8	BZ879331	BZ879331 CH240_195
32	36.8	5.4	457	2	BE679790	BE679790 df77d10.y
c 33	36.8	5.4	536	8	AZ450406	AZ450406 1M0249D04
34	36.8	5.4	661	4	BI385436	BI385436 BFL26_000
35	36.6	5.4	453	4	BI385638	BI385638 BFL26_000
36	36.6	5.4	705	2	BE357640	BE357640 DGI_21_B0
37	36.4	5.4	524	6	CB602899	CB602899 M3295_Mid
38	36.4	5.4	558	9	CG117372	CG117372 PUIGT77TB
39	36.4	5.4	596	6	CB390628	CB390628 OSTF138A2
40	36.4	5.4	930	9	CG041582	CG041582 PUJFE02TD
c 41	36.2	5.3	725	8	BZ260052	BZ260052 CH230-520
c 42	36.2	5.3	828	8	BZ248767	BZ248767 CH230-273
43	36	5.3	513	8	BH111996	BH111996 RPCI-24-2
44	36	5.3	645	9	CE190471	CE190471 tigr-gss-
45	36	5.3	679	8	BH188917	BH188917 037_I_05-

ALIGNMENTS

RESULT 1  
BI381780  
LOCUS  
DEFINITION  
BI381780 671 bp mRNA linear EST 26-AUG-2003  
BFLG1\_003411 Amphioxus 5-6 hrs cDNA library (Name convention: BFLG or MPMPGp498) Branchiostoma floridae cDNA clone MPMPGp498C047 5', mRNA sequence.  
ACCESSION  
BI381780  
VERSION  
BI381780.1 GI:30916875  
KEYWORDS  
EST.  
SOURCE  
Branchiostoma floridae (Florida lancelet)  
ORGANISM  
Branchiostoma floridae  
REFERENCE  
1 (bases 1 to 671)  
AUTHORS  
Panopoulou, G., Hennig, S., Groth, D., Krause, A., Poustka, A.J., Herwig, R., Vingron, M. and Lehrach, H.  
TITLE  
New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes  
JOURNAL  
Genome Res. 13 (6A), 1056-1066 (2003)  
MEDLINE  
22683279  
PUBMED  
12799346  
COMMENT  
Contact: Panopoulou G  
laboratory 145, dept. Lehrach  
Max-Planck-Institut fuer Molekulare Genetik  
Innestr. 63-73, D-14195 Berlin, Germany  
Tel: +49 30 8413 1235  
Fax: +49 30 8413 1128  
Email: panopoul@molgen.mpg.de  
The library was characterised by oligonucleotide fingerprinting (ONFP) to reduce sequencing redundancy. According to the ONFP procedure, clones giving the same hybridisation pattern with a battery of 200 8mer oligonucleotides are grouped into clusters. One clone per cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcript in the analysed library. The cluster size as well the coordinates of the rest of the clones assigned to the same fingerprint cluster as the clone from which the above EST is generated is available at the amphioxus project site at: <http://www.molgen.mpg.de/amphioxus/>  
Clones and filters are distributed via the Resource Center/Primary Database of the German Genome Project (<http://www.rzpd.de>)  
PCR Primers  
FORWARD: 5' CCCACGGCTTTACACTTTATGCTCCGGCTCG 3' (M13RSP)  
BACKWARD: 5' GCTATTACGCCAGCTGGCGAAGGGGATGTG 3' (M13FSP)  
Insert Length: 1200 Std Error: 0.00  
Seq primer: 5'-CCGGTCCGAATCCCGGT-3' pSport3/86  
High quality sequence stop: 671.





Db	444	TG	445	
RESULT 3				
BI379061				
LOCUS				
DEFINITION	BI379061 561 bp mRNA linear EST 26-AUG-2003			
	BFLG1 000692 Amphioxus 5-6 hrs cDNA library (Name convention: BFLG			
	or MPMGp498) Branchiostoma floridae cDNA clone MPMGp498P1084 5',			
	mRNA sequence.			
ACCESSION	BI379061			
VERSION	BI379061.1 GI:30914228			
KEYWORDS	EST.			
SOURCE	Branchiostoma floridae (Florida lancelet)			
ORGANISM	Branchiostoma floridae			
	Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;			
	Branchiostoma.			
REFERENCE	1 (bases 1 to 561)			
AUTHORS	Panopoulou,G., Hennig,S., Groth,D., Krause,A., Poustka,A.J.,			
	Herwig,R., Vingron,M. and Lehrach,H.			
TITLE	New evidence for genome-wide duplications at the origin of			
	vertebrates using an amphioxus gene set and completed animal			
	genomes			
JOURNAL	Genome Res. 13 (6A), 1056-1066 (2003)			
MEDLINE	22683279			
PUBMED	12799346			
COMMENT	Contact: Panopoulou G			
	laboratory 145, dept.Lehrach			
	Max-Planck-Institut fuer Molekulare Genetik			
	Innestr.63-73, D-14195 Berlin, Germany			
	Tel: +49 30 8413 1235			
	Fax: +49 30 8413 1128			
	Email: panopoul@molgen.mpg.de			
	The library was characterised by oligonucleotide fingerprinting			
	(ONFP) to reduce sequencing redundancy. According to the ONFP			
	procedure, clones giving the same hybridisation pattern with a			
	battery of 200 8mer oligonucleotides are grouped into clusters. One			
	clone per cluster is selected for sequencing. The size of each			
	cluster is an indicator of the frequency of a transcript in the			
	analysed library. The cluster size as well the coordinates of the			
	rest of the clones assigned to the same fingerprint cluster as the			
	clone from which the above EST is generated is available at the			
	amphioxus project site at: <a href="http://www.molgen.mpg.de/amphioxus/">http://www.molgen.mpg.de/amphioxus/</a>			
	Clones and filters are distributed via the Resource Center/Primary			
	Database of the German Genome Project ( <a href="http://www.rzpd.de">http://www.rzpd.de</a> )			
	PCR Primers			
	FORWARD: 5' CCCAGGCTTTACACTTTATGCTCCGGCTCG 3' (M13RSP)			
	BACKWARD: 5' GCTATTAGCCAGCTGGCGAAGGGGATGTG 3' (M13FSP)			
	Insert Length: 1200 Std Error: 0.00			
	Seq primer: 5'-CCGTCGCGAATTCCCGGT-3' pSport3/86			
	High quality sequence stop: 561.			
FEATURES	Location/Qualifiers			
source	1. .561			
	/organism="Branchiostoma floridae"			
	/mol_type="mRNA"			
	/db_xref="taxon:7739"			
	/clone="MPMGp498P1084"			
	/tissue_type="whole embryo"			
	/dev_stage="5-6 hrs (gastrula stage)"			
	/lab_host="E.coli, XLI blue"			
	/clone_lib="Amphioxus 5-6 hrs cDNA library (Name			
	convention: BFLG or MPMGp498)"			
	/note="Vector: pSport1; Site 1: SalI, KpnI, EcoRI (5');"			
	Site 2: NotI, BamHI, HindIII (3'); OligodT primed and			
	directionally cloned in pSport1 vector using a NotI			
	(5'-pGACTAGTTCTAGATCGGCGCGCC (T)15-3' and a SalI 5'-			
	TCGACCCACGGCTCCG-3'adapters (Gibco BRL)."			
ORIGIN				
	Query Match 7.1%; Score 48.4; DB 4; Length 561;			
	Best Local Similarity 47.8%; Pred. No. 0.0032;			
	Matches 173; Conservative 0; Mismatches 186; Indels 3; Gaps 1;			
QY	47	TTGCATGGAAGAACGGTCAATGGGCACGAGTTTGAATAGAGCGGCAAGGAGAGGGGA	106	
Db	97	TTACATCTTCGGCTCCGTCAACGGCGCGAGTTCGACCTGGTGGAGCGGAAAGGCA	156	
QY	107	GGCCATACGAAGGCCACAATACCGTAAAGCTTAAGGTAAACCAAGGGGGACCTTTGCCAT	166	
Db	157	ACCCGAACGATGG--AACGCTGGAGACCAGTGTGAAATCCACCCGGGCGCCCTGCCCT	213	
QY	167	TTGCTTGGGATATTTTGTCAACCAAAATTCAGTATGGAAGCAAGGTATATGTCAAGCACC	226	
Db	214	GCTCCCGCTGCTGATCGGACCCCAACCTGGGGTACGGCTTCTACCAGTACCTGCCCTTCC	273	
QY	227	CTGCCGACATACCAGACTATAAAAAGCTGTCTATTCTCTGAAGGATTAAATGGGAAAGG	286	
Db	274	CTGGCGGCGCTCACCCCTTCCAAACCGCCATCAGGACGGAGGGTACCAGGTTCAACGTG	333	
QY	287	TCATGAACCTTTGAAGACGGTGGCGTCTTACTGTAAACCCAGGATTCCAGTTTCAGGATG	346	
Db	334	TGTTCAAGTTGAAGACGGCGGAGTGTGAATTGCAACTTCCGCTACACATACGAGGGCG	393	
QY	347	GCTGTTTCATCTACAAGGTCAAGTTTCATTGGCGGTGAACCTTTCCTCCGATGGACCTGTTA	406	
Db	394	GCAAGATCAAAGGGGAGTTCACGCTGATAGGGTCAGGTTTCCCTGCCGCGGCCTGTGA	453	
QY	407	TG	408	
Db	454	TG	455	
RESULT 4				
BI386955				
LOCUS	BI386955 577 bp mRNA linear EST 26-AUG-2003			
DEFINITION	BFL26 001658 Amphioxus 26hr cDNA library (Name convention: BFL26 or			
	MPMGp531) Branchiostoma floridae cDNA clone MPMGp531M2297 5', mRNA			
	sequence.			
ACCESSION	BI386955			
VERSION	BI386955.1 GI:30921845			
KEYWORDS	EST.			
SOURCE	Branchiostoma floridae (Florida lancelet)			
ORGANISM	Branchiostoma floridae			
	Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;			
	Branchiostoma.			
REFERENCE	1 (bases 1 to 577)			
AUTHORS	Panopoulou,G., Hennig,S., Groth,D., Krause,A., Poustka,A.J.,			
	Herwig,R., Vingron,M. and Lehrach,H.			
TITLE	New evidence for genome-wide duplications at the origin of			
	vertebrates using an amphioxus gene set and completed animal			
	genomes			
JOURNAL	Genome Res. 13 (6A), 1056-1066 (2003)			
MEDLINE	22683279			
PUBMED	12799346			
COMMENT	Contact: Panopoulou G			
	laboratory 145, dept.Lehrach			
	Max-Planck-Institut fuer Molekulare Genetik			
	Innestr.63-73, D-14195 Berlin, Germany			
	Tel: +49 30 8413 1235			
	Fax: +49 30 8413 1128			
	Email: panopoul@molgen.mpg.de			
	The library was characterised by oligonucleotide fingerprinting			
	(ONFP) to reduce sequencing redundancy. According to the ONFP			
	procedure, clones giving the same hybridisation pattern with a			
	battery of 200 8mer oligonucleotides are grouped into clusters. One			
	clone per cluster is selected for sequencing. The size of each			
	cluster is an indicator of the frequency of a transcript in the			
	analysed library. The cluster size as well the coordinates of the			
	rest of the clones assigned to the same fingerprint cluster as the			
	clone from which the above EST is generated is available at the			
	amphioxus project site at: <a href="http://www.molgen.mpg.de/amphioxus/">http://www.molgen.mpg.de/amphioxus/</a>			
	Clones and filters are distributed via the Resource Center/Primary			
	Database of the German Genome Project ( <a href="http://www.rzpd.de">http://www.rzpd.de</a> )			
	PCR Primers			
	FORWARD: 5' CCCAGGCTTTACACTTTATGCTCCGGCTCG 3' (M13RSP)			
	BACKWARD: 5' GCTATTAGCCAGCTGGCGAAGGGGATGTG 3' (M13FSP)			
	Insert Length: 1200 Std Error: 0.00			
	Seq primer: 5'-CCGTCGCGAATTCCCGGT-3' pSport3/86			
	High quality sequence stop: 561.			
FEATURES	Location/Qualifiers			
source	1. .561			
	/organism="Branchiostoma floridae"			
	/mol_type="mRNA"			
	/db_xref="taxon:7739"			
	/clone="MPMGp498P1084"			
	/tissue_type="whole embryo"			
	/dev_stage="5-6 hrs (gastrula stage)"			
	/lab_host="E.coli, XLI blue"			
	/clone_lib="Amphioxus 5-6 hrs cDNA library (Name			
	convention: BFLG or MPMGp498)"			
	/note="Vector: pSport1; Site 1: SalI, KpnI, EcoRI (5');"			
	Site 2: NotI, BamHI, HindIII (3'); OligodT primed and			
	directionally cloned in pSport1 vector using a NotI			
	(5'-pGACTAGTTCTAGATCGGCGCGCC (T)15-3' and a SalI 5'-			
	TCGACCCACGGCTCCG-3'adapters (Gibco BRL)."			





QY	407	TG	408	
Db	441	TG	442	
RESULT 6				
BI382373	BI382373	591	bp	linear
LOCUS	BFLG2	000550	Amphioxus 5-6 hrs	cdNA library (Name convention: BFLG or MPMGp498) Branchiostoma floridae cdNA clone MPMGp498C1177 5', mRNA sequence.
DEFINITION	BI382373	1	GI:30917444	
ACCESSION	EST.			
VERSION	Branchiostoma floridae (Florida lancelet)			
KEYWORDS	Branchiostoma floridae			
SOURCE	Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae; Branchiostoma.			
ORGANISM	1 (bases 1 to 591)			
REFERENCE	Panopoulou,G., Hennig,S., Groth,D., Krause,A., Poustka,A.J., Herwig,R., Vingron,M. and Lehrach,H.			
AUTHORS	New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes			
TITLE	Genomes			
JOURNAL	Genome Res. 13 (6A), 1056-1066 (2003)			
MEDLINE	22683279			
PUBMED	12799346			
COMMENT	Contact: Panopoulou G Laboratory 145, dept.Lehrach Max-Planck-Institut fuer Molekulare Genetik Innestr.63-73, D-14195 Berlin, Germany Tel: +49 30 8413 1235 Fax: +49 30 8413 1128 Email: panopoul@molgen.mpg.de The library was characterised by oligonucleotide fingerprinting (ONFP) to reduce sequencing redundancy. According to the ONFP procedure, clones giving the same hybridisation pattern with a battery of 200 8mer oligonucleotides are grouped into clusters. One clone per cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcript in the analysed library. The cluster size as well the coordinates of the rest of the clones assigned to the same fingerprint cluster as the clone from which the above EST is generated is available at the amphioxus project site at: <a href="http://www.molgen.mpg.de/amphioxus/">http://www.molgen.mpg.de/amphioxus/</a> Clones and filters are distributed via the Resource Center/Primary Database of the German Genome Project ( <a href="http://www.rzpd.de">http://www.rzpd.de</a> ) PCR Primers FORWARD: 5' CCCAGGCTTTACACTTATGCTTCGGCTCG 3' (M13RSP) BACKWARD: 5' GCTATTACGCCAGCTGGCGAAGGGGATGTG 3' (M13FSP) Insert Length: 1 Std Error: 0.00 Seq primer: 5'-CCGGTCCGGAATTCGCGGT-3' pSport3/86 High quality sequence stop: 591. Location/Qualifiers 1. .591 /organism="Branchiostoma floridae" /mol_type="mRNA" /db_xref="taxon:7739" /clone="MPMGp498C1177" /tissue_type="whole embryo" /dev_stage="5-6 hrs (gastrula stage)" /lab_host="E.coli, XLI blue" /clone_lib="Amphioxus 5-6 hrs cdNA library (Name convention: BFLG or MPMGp498)" /note="Vector: pSport1; Site_1: SalI, KpnI, EcoRI (5'); Site_2: NotI, BamHI, HindIII (3'); OligodT primed and directionally cloned in pSport1 vector using a NotI (5'-pGACTAGTTCTAGATCGGAGCGGCCCC (T)15-3' and a SalI 5'-TCGACCCACGCGTCCG-3'adapters (Gibco BRL)."			
ORIGIN				
Query Match	6.7%	Score	45.2;	DB 4; Length 591;
Best Local Similarity	47.2%	Pred. No.	0.032;	
Matches	171;	Conservative	0;	Mismatches 188; Indels 3; Gaps 1;

QY	47	TTCCGATCGAAGGAACGGTCAATGGCAGCAGTTTGAATAAGAGCGCAAGAGAGGGGA	106
Db	96	TTACATCTTCGGCTCCGTCAATGGCGCGAGTTCGACCTGGTGGAGCGGAAAGGGCA	155
QY	107	GGCCATACGAAGGCCACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCAT	166
Db	156	ACCCGAATGATGG--AACGCTGGAGACCAAGTGTGAAATCCACCCGGGGCGCTGCCCT	212
QY	167	TTGCTTGGGATATTTTGTCAACCAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACC	226
Db	213	GCTCCCGCTGTGTGATCGGACCCCAACCTGGGGTACGGCTTCTACCACTACCTGCCCTTC	272
QY	227	CTGCCGACATACCACTATAAAAAGCTGTTCATTTCTTGAAGGATTTAAATGGGAAAGGG	286
Db	273	CTGGCGGCGCTCACCCCTTCCAAACCGCCATCACGGACGGAGGTACCAAGTTTCAACCGTG	332
QY	287	TCATGAACCTTTGAAGACGGTGGCGTCTTACTGTAAACCCAGGATTCAGTTTGCAGGATG	346
Db	333	TGTTCAAGTTTGAAGACGGAGGATTGCTGAATTGCAACTTCCGCTACACCTACGAGGGCG	392
QY	347	GCTGTTTCATCTACAAGGTCAAGTTCAATTGGCGTGAACTTTCCTTCCGATGGACCTGTTA	406
Db	393	GCAAGATCAAAGGGGAGTTCCAGCTGATCGGGTCAGGTTTCCCTGCGCGCGGCTGTGA	452
QY	407	TG	408
Db	453	TG	454
RESULT 7			
BI382442	BI382442	707	bp
LOCUS	BFLG2	000619	Amphioxus 5-6 hrs
DEFINITION	or MPMGp498) Branchiostoma floridae cdNA clone MPMGp498C08131 5', mRNA sequence.		
ACCESSION	BI382442		
VERSION	BI382442.1	GI:30917513	
KEYWORDS	EST.		
SOURCE	Branchiostoma floridae (Florida lancelet)		
ORGANISM	Branchiostoma floridae		
REFERENCE	Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae; Branchiostoma.		
AUTHORS	1 (bases 1 to 707)		
TITLE	Panopoulou,G., Hennig,S., Groth,D., Krause,A., Poustka,A.J., Herwig,R., Vingron,M. and Lehrach,H.		
JOURNAL	New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes		
MEDLINE	Genome Res. 13 (6A), 1056-1066 (2003)		
PUBMED	22683279		
COMMENT	12799346		
Contact: Panopoulou G			
Laboratory 145, dept.Lehrach			
Max-Planck-Institut fuer Molekulare Genetik			
Innestr.63-73, D-14195 Berlin, Germany			
Tel: +49 30 8413 1235			
Fax: +49 30 8413 1128			
Email: panopoul@molgen.mpg.de			
The library was characterised by oligonucleotide fingerprinting (ONFP) to reduce sequencing redundancy. According to the ONFP procedure, clones giving the same hybridisation pattern with a battery of 200 8mer oligonucleotides are grouped into clusters. One clone per cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcript in the analysed library. The cluster size as well the coordinates of the rest of the clones assigned to the same fingerprint cluster as the clone from which the above EST is generated is available at the amphioxus project site at: <a href="http://www.molgen.mpg.de/amphioxus/">http://www.molgen.mpg.de/amphioxus/</a> Clones and filters are distributed via the Resource Center/Primary Database of the German Genome Project ( <a href="http://www.rzpd.de">http://www.rzpd.de</a> ) PCR Primers FORWARD: 5' CCCAGGCTTTACACTTATGCTTCGGCTCG 3' (M13RSP) BACKWARD: 5' GCTATTACGCCAGCTGGCGAAGGGGATGTG 3' (M13FSP) Insert Length: 1 Std Error: 0.00 Seq primer: 5'-CCGGTCCGGAATTCGCGGT-3' pSport3/86 High quality sequence stop: 591. Location/Qualifiers 1. .591 /organism="Branchiostoma floridae" /mol_type="mRNA" /db_xref="taxon:7739" /clone="MPMGp498C1177" /tissue_type="whole embryo" /dev_stage="5-6 hrs (gastrula stage)" /lab_host="E.coli, XLI blue" /clone_lib="Amphioxus 5-6 hrs cdNA library (Name convention: BFLG or MPMGp498)" /note="Vector: pSport1; Site_1: SalI, KpnI, EcoRI (5'); Site_2: NotI, BamHI, HindIII (3'); OligodT primed and directionally cloned in pSport1 vector using a NotI (5'-pGACTAGTTCTAGATCGGAGCGGCCCC (T)15-3' and a SalI 5'-TCGACCCACGCGTCCG-3'adapters (Gibco BRL)."			
ORIGIN			
Query Match	6.7%	Score	45.2;
Best Local Similarity	47.2%	Pred. No.	0.032;
Matches	171;	Conservative	0;
		Mismatches	188;
		Indels	3;
		Gaps	1;





Db 376 GCAAGATCAAAGGGAGTTCCACCTGCTCGGTGTCAGGTTTCCCTGCCGGCGGCGCTGTGA 435

QY 407 TG 408

Db 436 TG 437

RESULT 9

BI384961

LOCUS

DEFINITION

BFLG2\_003138 Amphioxus 5-6 hrs cDNA library (Name convention: BFLG or MPMGp498) Branchiostoma floridae cDNA clone MPMGp498E05107 5', mRNA sequence.

ACCESSION

BI384961

VERSION

BI384961.1 GI:30919957

KEYWORDS

EST.

SOURCE

Branchiostoma floridae (Florida lancelet)

ORGANISM

Branchiostoma floridae

Branchiostoma floridae

Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae; Branchiostoma.

REFERENCE

1 (bases 1 to 629)

AUTHORS

Panopoulou,G., Hennig,S., Groth,D., Krause,A., Poustka,A.J., Herwig,R., Vingron,M. and Lehrach,H.

TITLE

New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes

JOURNAL

Genome Res. 13 (6A), 1056-1066 (2003)

MEDLINE

22683279

PUBMED

12799346

COMMENT

Contact: Panopoulou G

Laboratory 145, dept.Lehrach

Max-Planck-Institut fuer Molekulare Genetik

Imneistr.63-73, D-14195 Berlin, Germany

Tel: +49 30 8413 1235

Fax: +49 30 8413 1128

Email: panopoul@molgen.mpg.de

The library was characterised by oligonucleotide fingerprinting (ONFP) to reduce sequencing redundancy. According to the ONFP procedure, clones giving the same hybridisation pattern with a battery of 200 8mer oligonucleotides are grouped into clusters. One clone per cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcript in the analysed library. The cluster size as well the coordinates of the rest of the clones assigned to the same fingerprint cluster as the clone from which the above EST is generated is available at the amphioxus project site at: <http://www.molgen.mpg.de/amphioxus/>

Clones and filters are distributed via the Resource Center/Primary Database of the German Genome Project (<http://www.rzpd.de>)

PCR Primers

FORWARD: 5' CCCAGGCTTTACACTTTATGCTCCGGCTCG 3' (M13RSP)

BACKWARD: 5' GCTATTACGCCAGCTGGCGAAAGGGGATGTG 3' (M13FSP)

Insert Length: 1 Std Error: 0.00

Seq primer: 5'-CCGTCGGGAATTCGCGGT-3' pSPORT3/86

High quality sequence stop: 629.

Location/Qualifiers

1. .629

/organism="Branchiostoma floridae"

/mol\_type="mRNA"

/db\_xref="taxon:7739"

/clone="MPMGp498E05107"

/tissue\_type="whole embryo"

/dev\_stage="5-6 hrs (gastrula stage)"

/lab\_host="E.coli, XLI blue"

/clone\_lib="Amphioxus 5-6 hrs cDNA library (Name convention: BFLG or MPMGp498)"

/note="Vector: pSPORT1; Site 1: SalI, KpnI, EcoRI (5'); Site 2: NotI, BamHI, HindIII (3'); OligodT primed and directionally cloned in pSPORT1 vector using a NotI (5'-pGACTAGTTCTAGATCGGAGCGCGCCC (T)15-3' and a SalI 5'-TCGACCCACGCTCCG-3' adapters (Gibco BRL)."

ORIGIN

Query Match 6.4%; Score 43.2; DB 4; Length 629; .

Best Local Similarity 46.2%; Pred. No. 0.13;

Matches 181; Conservative 0; Mismatches 208; Indels 3; Gaps 1;

QY 47 TTCGCATGAAGGAACGGTCAATGGCAGCAGAGTTTGAATAAGAGGCAAGGAGAGGGGA 106

Db 112 TGCACATCTATGGTCCATCAACGGTATGGAGTTCGACATGGTGGCGGTGGAGTGGA 171

QY 107 GGCCATACGAAGGCCACAATACCGTAAAGCTTAAGGTAAACCAAGGGGGACCTTTGCCAT 166

Db 172 ACCCAAAGGACGGATCGCTGAGCGTAAACATGAAGTCTACCAAAGGTGCCC---TGCGCG 228

QY 167 TTGCTTGGGATATTTTGTCAACCAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACC 226

Db 229 TTTCTCTCTGCTGGTGGGCCCGCATCTGGGGTACGGCCACTATCATGACCTACCTTCC 288

QY 227 CTGCCGACATACCAGACTATATAAAAGCTGTCTATTTCTCTGAAGGATTTAAATGGAAAGG 286

Db 289 CTGACGGTCCGTGCGCTTTCCAGGCGAGCTGTGAACAACGGCGGTATCAAAATGCATCGCT 348

QY 287 TCATGAACTTTGAAGACGGTGGCTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATG 346

Db 349 CTTTCAACTTCGAGGACGGTGCCGTGCTGACTGCCACCTACAATACTCTCTACAGCGCG 408

QY 347 GCTGTTTCATCTACAAGGTCAAGTTCATTTGCGGTGAACTTTCTCTCCGATGGACCTGTTA 406

Db 409 GCAAGATCCATGGAGAGTTTTCATCTGTGGGCAGCTGTTTCCCCGACGATAGTCCGGTGA 468

QY 407 TGCAAAAGAAGACAATGGGCTGGGAAGCCAGC 438

Db 469 TGACCAACGGCTGACCGTTTTTGGACAGGAGC 500

RESULT 10

CL065858/c

LOCUS

DEFINITION

CL065858

CL065858.1 GI:40521771

ACCESSION

CL065858

VERSION

GSS.

KEYWORDS

SOURCE

ORGANISM

Xenopus tropicalis (western clawed frog)

Xenopus tropicalis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae; Xenopus; Silurana.

REFERENCE

1 (bases 1 to 930)

AUTHORS

Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T., Mardis,E. and Wilson,R.

TITLE

A physical map of the xenopus tropicalis genome

JOURNAL

Unpublished (2003)

COMMENT

Contact: Richard K Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: [submissions@watson.wustl.edu](mailto:submissions@watson.wustl.edu)

Insert Length: 175000 Std Error: 0.00

Seq primer: RM1 TACGACTCACTATAGGGAGA

Class: BAC ends

High quality sequence start: 3

High quality sequence stop: 510.

Location/Qualifiers

1. .930

/organism="Xenopus tropicalis"

/mol\_type="genomic DNA"

/strain="Nigerian frog"

/db\_xref="taxon:8364"

/clone="CH216-105J9"

/sex="male"

/cell\_line="Stock 248 F7A2, inbred N7"

/clone\_lib="CH216"

/note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis BAC library"

ORIGIN









ORIGIN

Query Match  
Best Local Similarity 5.8%; Score 39; DB 8; Length 808;  
Matches 72; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 434 CCAGCACTGAGCGTTTGTATCCTCGTGATGGCGGTGAAAGGAGAGATTCATAAGGCTC 493  
|||||  
Db 256 CCAGCACTGTCTGCTGCATGCACCAAGCTTCCTGCCATGATGATAATAGACTAAGCCTC 315  
  
QY 494 TGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAGTATTTACATGGCAAAGA 553  
|||||  
Db 316 TGAATTGTAAGCCAGCCCCCAATTAAATGTTTCTTTATAAGAGTTTCCATGGTTAAGG 375  
  
QY 554 AGCCTGT 560  
|||||  
Db 376 TGCCTCT 382

ORIGIN

Query Match  
Best Local Similarity 5.7%; Score 38.6; DB 5; Length 571;  
Matches 74; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 174 GGATATTTTGTCAACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCTGCCGA 233  
|||||  
Db 211 GGATATGCTCAATCCACAATTTTGAAGAAATAGGCAAGGAATTTGTTAATCACTACTTTCA 270  
  
QY 234 CATACCAGACTATAAAAAGCTGTCAATTCCTGAAGGATTTTAAATGGGAAAGGGTCATGAA 293  
|||||  
Db 271 GTTATTTAACAGTGGAGGAATGAATTTGGCTGCATTATATAAAGATATAAGTATGATGAG 330  
  
QY 294 CTTTGAAGACGGT 306  
|||||  
Db 331 TTTTGAATAATGAT 343

Search completed: July 1, 2005, 17:24:20  
Job time : 3048 secs

RESULT 15

BQ451564

LOCUS

DEFINITION

PFESToab06b02.y1 Plasmodium falciparum 3D7 asexual cDNA Plasmodium falciparum 3D7 cDNA 5' similar to SW:NTF2\_NEUCR P87102 NUCLEAR TRANSPORT FACTOR 2 ;, mRNA sequence.

ACCESSION

BQ451564

VERSION

BQ451564.1 GI:21254676

KEYWORDS

EST.

SOURCE

Plasmodium falciparum 3D7

ORGANISM

Plasmodium falciparum 3D7

REFERENCE

1 (bases 1 to 571)

AUTHORS

Tang,K., Cole,R., Chakrabarti,D., Haywood,R., Clifton,S., Pape,D., Marra,M., Hillier,L., Martin,J., Wylie,T., Dante,M., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Jentes,E., Ronko,I., Tsagareishvili,R., Belaygorod,L., Franklin,C., Carr,L., Grow,A., Maguire,L., Richey,J., Wadkins,J., Kennedy,S., Levinso,D., Waterston,R., Wilson,R. and Sibley,D.

TITLE

WashU Plasmodium EST Project

JOURNAL

Unpublished (2001)

COMMENT

Contact: L. David Sibley  
WashU Plasmodium EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
Library was constructed by Debopam Chakrabarti DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: L. David Sibley (sibley@borcim.wustl.edu), Washington University  
Seq primer: -40UP from Gibco  
High quality sequence stop: 420.

FEATURES

source

1..571

/organism="Plasmodium falciparum 3D7"

/mol\_type="mRNA"

/db\_xref="taxon:36329"

/lab\_host="DH10B (GeneHog, Invitrogen, Inc.)"

/clone\_lib="plasmodium falciparum 3D7 asexual cDNA"

/note="vector: pBluescript SK plus; Site 1: EcoRI; Site 2: XhoI; Library was constructed by Debopam Chakrabarti. Total RNA samples were isolated from mixed stage saponin(0.1%-lysed P. falciparum 3D7 infected erythrocytes by the acidic guanidinium-phenol chloroform method. The poly A+ RNA was isolated by the polyAT-Tract mRNA isolation system (Promega, WI) using streptavidin Magnosphere particles. Directional cDNA libraries were constructed by oligo d(T) priming of poly(A)+ RNA (5mg) into EcoRI and XhoI sites of l ZapII vector using the Zap cDNA synthesis kit (Stratagene, CA). The average size of the cDNA inserts in the library was between 1.0 and 1.5kb. Clones were mass excised using the ExAssist helper phage

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2005, 12:17:58 ; Search time 1029 Seconds  
(without alignments)  
3900.471 Million cell updates/sec

Title: US-10-006-922A-11  
Perfect score: 678  
Sequence: 1 atgagggtcttccaagaatgt.....gccaccatctgttcctttaa 678

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score		Query Match		Length	DB ID	Description
1	678	100.0	678	4	AAH47654	AAH47654	Discosoma
2	678	100.0	678	4	AAD13053	AAD13053	Discosoma
3	678	100.0	678	4	AAD11142	AAD11142	Discosoma
4	678	100.0	678	6	ABA95921	ABA95921	Yeast opt
5	678	100.0	678	10	ADC24127	ADC24127	Discosoma
6	678	100.0	678	11	ADL46204	ADL46204	Discosoma
7	678	100.0	678	11	ADN33978	ADN33978	Wild-type
8	678	100.0	678	12	ADI36420	ADI36420	Discosoma
9	678	100.0	678	12	ADM97768	ADM97768	D sp red
10	678	100.0	859	6	AAL47952	AAL47952	Discosoma
11	678	100.0	859	8	AAD53432	AAD53432	Discosoma
12	678	100.0	859	10	AAD61969	AAD61969	Discosoma
13	677	99.9	678	10	ADF70404	ADF70404	Discosoma
14	677	99.9	3311	9	ACA62995	ACA62995	Plasmid D
15	665	98.1	666	6	ABL61142	ABL61142	Red fluor
16	661	97.5	711	12	ADO78067	ADO78067	Corallimo
17	661	97.5	711	12	ADO78069	ADO78069	Corallimo
18	658.8	97.2	898	4	AAD03615	AAD03615	Discosoma
19	657.2	96.9	898	3	AAA48742	AAA48742	Discosoma
20	642.8	94.8	678	12	ADM13536	ADM13536	DNA encod

21	566.4	83.5	876	3	AAA28958	AAA28958	Discosoma
22	566.4	83.5	876	4	AAD03618	AAD03618	Discosoma
23	566.4	83.5	876	10	ACC44649	ACC44649	Red fluor
24	450.6	66.5	723	8	ABZ22476	ABZ22476	Mammalian
25	449.8	66.3	681	4	AAH47656	AAH47656	Anthozoan
26	449.2	66.3	678	6	AAD46278	AAD46278	Discosoma
27	449.2	66.3	678	6	AAD28207	AAD28207	Discosoma
28	449.2	66.3	695	3	AAA48743	AAA48743	Humanised
29	447.6	66.0	678	6	ABA95922	ABA95922	Yeast opt
30	447.2	66.0	681	10	ADC24134	ADC24134	Discosoma
31	447.2	66.0	681	11	ADL46205	ADL46205	Human cod
32	447.2	66.0	681	11	ADL46225	ADL46225	Human cod
33	447.2	66.0	2721	13	ADS75466	ADS75466	Fibrohexa
34	447.2	66.0	2772	13	ADS75468	ADS75468	Fibrohexa
35	447.2	66.0	4692	6	AAL47954	AAL47954	Modified
36	447.2	66.0	4692	10	ACC44640	ACC44640	Vector pd
37	447.2	66.0	7910	4	AAD09979	AAD09979	pBit(dHSP
38	447.2	66.0	9320	6	ABS56664	ABS56664	Plasmid p
39	447.2	66.0	9658	12	ADI04104	ADI04104	Vector pt
40	447.2	66.0	9678	12	ADI04103	ADI04103	Vector pt
41	447.2	66.0	10263	12	ADI04102	ADI04102	Vector pt
42	446	65.8	678	6	AAD28208	AAD28208	Discosoma
43	445.8	65.8	1638	10	ADL18131	ADL18131	RFP:PS(NI
44	445.8	65.8	1647	10	ADL18155	ADL18155	RFP:PS(HI
45	445.8	65.8	6893	10	ADE24111	ADE24111	Proviral

ALIGNMENTS

RESULT 1  
AAH47654  
ID AAH47654 standard; cDNA; 678 BP.

AC AAH47654;

XX 30-NOV-2001 (first entry)

XX Discosoma sp. red drFP583 protein coding sequence.

XX Fluorescent protein; Anthozoan; fluorescence; marker; FRET; drFP583; ss.

XX Discosoma sp.

XX WO200162919-A1.

XX 30-AUG-2001.

XX 13-FEB-2001; 2001WO-US004625.

XX 23-FEB-2000; 2000US-0184732P.

XX (AURO-) AURORA BIOSCIENCES CORP.

XX Nelson D, Zamaira E, Tsien R;

XX WPI; 2001-557704/62.

XX Proteins for Fluorescence Resonance Energy Transfer (FRET) comprise functional red fluorescent proteins, and the encoding nucleic acids, with key mutations for improving the proteins function.

XX Disclosure; Page 83; 90pp; English.

XX The invention provides a nucleic acid encoding functional red fluorescent CC protein (II) that differs from the sequence of an Anthozoan red fluorescent protein by at least one amino acid substitution, and with different fluorescent properties. The red fluorescent protein of the invention can be expressed by standard recombinant methodology. (II) are used a fluorescent markers and FRET partners. It is used for identifying protein-protein interactions. (II) is also suitable for multiplexed fluorescent analysis and FRET-based applications using existing Aequorea fluorescent proteins. (II) has improved brightness, reduced spectral

Walden

CC cross talk, and is rapidly and efficiently expressed in mammalian cells.  
CC The key mutations in the encoding nucleic acids provide improved folding,  
CC brightness, and create (II) with sharper, more defined excitation and  
CC emission peaks when expressed in mammalian cells. The present sequence  
CC represents the coding sequence of a *Discosoma* sp. drFP583 protein, an  
CC anthozoan fluorescent protein  
XX  
SQ Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Query Match 100.0%; Score 678; DB 4; Length 678;  
Best Local Similarity 100.0%; Pred. No. 9.8e-208;  
Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGGTCCTCCAAGAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTGCGCATGGAAGGA 60  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1 ATGAGGTCCTCCAAGAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTGCGCATGGAAGGA 60  
Qy 61 ACGGTCAATGGGCACGAGTTTGAATAGAACGCGAAGGAGAGGGGAGGCCATACGAAGGC 120  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
61 ACGGTCAATGGGCACGAGTTTGAATAGAACGCGAAGGAGAGGGGAGGCCATACGAAGGC 120  
Qy 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
Qy 181 TTGTCACCAACAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCTGCCGACATACCA 240  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
181 TTGTCACCAACAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCTGCCGACATACCA 240  
Qy 241 GACTATAAAAGCTGTCTATTCTGGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAA 300  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
241 GACTATAAAAGCTGTCTATTCTGGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAA 300  
Qy 301 GACGGTGGCGTCTACTGTAAACCCAGGATTCAGTTTGACGATGGCTGTTTCATCTAC 360  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
301 GACGGTGGCGTCTACTGTAAACCCAGGATTCAGTTTGACGATGGCTGTTTCATCTAC 360  
Qy 361 AAGGTCAAGTTCAATGGCGTGAACTTTCCCTCCGATGGACCTGTTATGCAAAAGAGACA 420  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
361 AAGGTCAAGTTCAATGGCGTGAACTTTCCCTCCGATGGACCTGTTATGCAAAAGAGACA 420  
Qy 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
Qy 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT 540  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT 540  
Qy 541 TACATGGCAAGAAGCCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAT 600  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
541 TACATGGCAAGAAGCCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAT 600  
Qy 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGGAGCGC 660  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGGAGCGC 660  
Qy 661 CACCATCTGTTCTTTAA 678  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
661 CACCATCTGTTCTTTAA 678

RESULT 2  
AAD13053  
ID AAD13053 standard; cDNA; 678 BP.  
XX  
AC AAD13053;  
XX  
DT 16-OCT-2001 (first entry)  
XX  
DE *Discosoma* sp. "red" anthozoa fluorescent protein, drFP583 cDNA.  
XX  
KW Protein destabilisation; linker moiety; reporter moiety; disease model;

KW linear multimerised domain; -NH-ubiquitin protein endoprotease;  
KW transgenic animal; transgenic plant; disease resistance;  
KW anthozoa fluorescent protein; natural fluorescent protein; ss.  
XX  
OS *Discosoma* sp.  
XX  
PN WO200157242-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 02-FEB-2001; 2001WO-US003791.  
XX  
PR 04-FEB-2000; 2000US-00498098.  
XX  
PA (AURO-) AURORA BIOSCIENCES CORP.  
XX  
PI Stack JH, Whitney M, Cubitt AB, Pollok BA;  
XX  
DR WPI; 2001-488890/53.  
XX

Destabilizing proteins in living cells, by coupling a target protein to  
linear multimerized destabilization domain non-cleavable by -NH-ubiquitin  
protein endoproteases, comprising two copies of the domain.  
XX  
PS Disclosure; Page 110; 171pp; English.

XX  
CC The present invention relates to a method for destabilising a target  
CC protein in a cell. The method comprises a linker moiety which operatively  
CC couples a target protein (a reporter moiety) to a linear multimerised  
CC destabilising domain, which is non-cleavable by a -NH-ubiquitin protein  
CC endoproteases. The method is useful for detecting an activity such as  
CC protease, protein kinase or phosphoprotein phosphatase activity and is  
CC also useful for identifying modulators of these activities. The method is  
CC also useful for developing novel assays for a wide range of post-  
CC translational activities, such as proteolysis, phosphorylation,  
CC dephosphorylation, glycosylation, methylation, sulfation, prenylation,  
CC disulfide bond formation and ADP-ribosylation within cells. The  
CC recombinant DNA molecule of the invention is useful for creating  
CC transgenic animals useful as disease models and transgenic plants with  
CC improved disease resistance or other favourable traits. The present  
CC sequence is *Discosoma* sp. "red" anthozoa fluorescent protein, drFP483  
CC cDNA which is a natural fluorescent protein used as a reporter moiety in  
CC the exemplification of the invention  
XX

SQ Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Query Match 100.0%; Score 678; DB 4; Length 678;  
Best Local Similarity 100.0%; Pred. No. 9.8e-208;  
Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGGTCCTCCAAGAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTGCGCATGGAAGGA 60  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1 ATGAGGTCCTCCAAGAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTGCGCATGGAAGGA 60  
Qy 61 ACGGTCAATGGGCACGAGTTTGAATAGAACGCGAAGGAGAGGGGAGGCCATACGAAGGC 120  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
61 ACGGTCAATGGGCACGAGTTTGAATAGAACGCGAAGGAGAGGGGAGGCCATACGAAGGC 120  
Qy 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
Qy 181 TTGTCACCAACAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCTGCCGACATACCA 240  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
181 TTGTCACCAACAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCTGCCGACATACCA 240  
Qy 241 GACTATAAAAGCTGTCTATTCTGGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAA 300  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
241 GACTATAAAAGCTGTCTATTCTGGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAA 300  
Qy 301 GACGGTGGCGTCTACTGTAAACCCAGGATTCAGTTTGACGATGGCTGTTTCATCTAC 360  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
301 GACGGTGGCGTCTACTGTAAACCCAGGATTCAGTTTGACGATGGCTGTTTCATCTAC 360

QY 361 AAGGTCAAGTTCATTGGCGTGAACCTTTCTTCGATGGACCTGTTATGCAAAAGAAGACA 420  
Dbb 361 AAGGTCAAGTTCATTGGCGTGAACCTTTCTTCGATGGACCTGTTATGCAAAAGAAGACA 420  
QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCTCGTGATGGCGTGTGAAAGGAGAG 480  
Dbb 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCTCGTGATGGCGTGTGAAAGGAGAG 480  
QY 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTGCATTACCTAGTTGAATTCAAAAAGTATT 540  
Dbb 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTGCATTACCTAGTTGAATTCAAAAAGTATT 540  
QY 541 TACATGGCAAAAGACCTCTGTCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600  
Dbb 541 TACATGGCAAAAGACCTCTGTCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600  
QY 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC 660  
Dbb 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC 660  
QY 661 CACCATCTGTTCTCTTTAA 678  
Dbb 661 CACCATCTGTTCTCTTTAA 678

RESULT 3

AAD111142  
ID AAD111142 standard; DNA; 678 BP.  
XX AAD111142;  
AC AAD111142;  
XX  
DT 24-SEP-2001 (first entry)  
DE Discosoma species Anthozoa fluorescent protein, drFP583 encoding DNA.  
XX  
KW Transmembrane potential; biological membrane; fluorescent ion; detection;  
KW test chemical screening; Anthozoa fluorescent protein; FP;  
KW transgenic organism; drFP583 protein; ds.  
XX  
OS Discosoma sp.  
XX WO200142211-A2.  
PN  
PD 14-JUN-2001.  
XX  
PF 12-DEC-2000; 2000WO-US033739.  
XX  
PR 13-DEC-1999; 99US-00459956.  
XX (REGC ) UNIV CALIFORNIA.  
PA  
XX Tsien RY, Gonzalez JE;  
PI WPI; 2001-457276/49.  
XX  
PT Determining electrical potential across a membrane in biological systems,  
PT comprises introducing two reagents, exposing the membrane to light and  
PT measuring the energy transfer.  
XX  
PS Disclosure; Page 150; 154pp; English.

CC The patent discloses optical methods and compositions for determining  
CC transmembrane potential across biological membranes in living cells. The  
CC method of determining the electrical potential across a membrane in a  
CC biological system comprises introduction of two reagents, a first reagent  
CC comprising a hydrophobic fluorescent ion capable of redistributing from a  
CC first face of the membrane to a second face of the membrane in response  
CC to membrane potential change and a second reagent that label the first  
CC face or the second face of the membrane, which comprises a chromophore  
CC capable of undergoing energy transfer by either donating or accepting  
CC excited state energy to the fluorescent ion. The membrane is then exposed  
CC to excitation light and the energy transfer between the reagents is

CC measured and related to the membrane potential. The method is useful for  
CC detecting changes in membrane potential in subcellular organelle  
CC membranes in biological systems. The method is used for screening of test  
CC chemicals for activity to modulate the activity of target ion channel.  
CC The invention also provides a transgenic organism comprising a first  
CC reagent that comprises a charged hydrophobic fluorescent molecule and a  
CC second reagent comprising a bioluminescent or naturally fluorescent  
CC protein. The present sequence is Discosoma species (red) DNA encoding an  
CC Anthozoa fluorescent protein (FP), drFP583  
XX

SQ Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Query Match 100.0%; Score 678; DB 4; Length 678;

Best Local Similarity 100.0%; Pred. No. 9.8e-208;

Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTCATGAGGTTTAAAGTTCGATGGAAGGA 60  
Dbb 1 ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTCATGAGGTTTAAAGTTCGATGGAAGGA 60  
QY 61 ACGGTCAATGGGCACGAGTTTGAATAGAACGCGAAGGAGAGGGAGGCACATACGAAGGC 120  
Dbb 61 ACGGTCAATGGGCACGAGTTTGAATAGAACGCGAAGGAGAGGGAGGCACATACGAAGGC 120  
QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
Dbb 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
QY 181 TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCTTGCCGACATACCA 240  
Dbb 181 TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCTTGCCGACATACCA 240  
QY 241 GACTATAAAAAGCTGTCTATTTCTGAAGGATTTAAATGGAAAGGGTCATGAACCTTTGAA 300  
Dbb 241 GACTATAAAAAGCTGTCTATTTCTGAAGGATTTAAATGGAAAGGGTCATGAACCTTTGAA 300  
QY 301 GACGGTGGCGTCGTACTGTAAACCCAGGATTCAGTTTGACAGGATGGCTGTTTCATCTAC 360  
Dbb 301 GACGGTGGCGTCGTACTGTAAACCCAGGATTCAGTTTGACAGGATGGCTGTTTCATCTAC 360  
QY 361 AAGGTCAAGTTCATTGGCGTGAACCTTTCTTCGATGGACCTGTTATGCAAAAGAAGACA 420  
Dbb 361 AAGGTCAAGTTCATTGGCGTGAACCTTTCTTCGATGGACCTGTTATGCAAAAGAAGACA 420  
QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
Dbb 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
QY 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCACTACCTAGTTGAATTCAAAAAGTATT 540  
Dbb 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCACTACCTAGTTGAATTCAAAAAGTATT 540  
QY 541 TACATGGCAAAAGACCTCTGTCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600  
Dbb 541 TACATGGCAAAAGACCTCTGTCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600  
QY 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC 660  
Dbb 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC 660  
QY 661 CACCATCTGTTCTCTTTAA 678  
Dbb 661 CACCATCTGTTCTCTTTAA 678

RESULT 4

ABA95921  
ID ABA95921 standard; DNA; 678 BP.  
XX  
AC ABA95921;  
XX  
DT 29-MAY-2002 (first entry)  
XX





CC in reduced propensity of the DsRed variant to form tetramers. (I) is  
CC useful for detecting transcriptional activity by providing a host cells  
CC containing a vector which comprises (I) operatively linked to an  
CC expression control sequence, and an unit to assay the variant fluorescent  
CC protein fluorescence, and assaying fluorescence of the variant  
CC fluorescent protein produced by (VII), where variant fluorescent protein  
CC fluorescence is indicative of transcriptional activity. A polynucleotide  
CC encoding a fusion protein is useful for the analysis of in vivo  
CC localisation or trafficking of a polypeptide of interest. A polypeptide  
CC marker is useful as markers to identify the location and amount of a  
CC target protein produced, where the target protein is fused to the marker,  
CC as a complement to or alternative for the green fluorescent protein or  
CC its spectral variant, for detecting induction of transcription, in  
CC applications involving fluorescence energy resonance transfer (FRET),  
CC which detects events as the function of the movement of fluorescent  
CC donors and acceptors towards or away from each other, for making  
CC fluorescent sensors for protein kinase and phosphatase activities or  
CC indicators for ions and molecules such as Ca2+, Zn2+, for identifying the  
CC presence of a molecule in a sample, for identifying a specific  
CC interaction of a first and second molecule, for determining whether a  
CC sample contains an enzyme or for determining the pH of the sample. (I) is  
CC useful for identifying a region or condition that regulates the activity  
CC of an expression control sequence. This sequence encodes Discosoma wild-  
CC type red fluorescent protein.

SQ Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Query Match 100.0%; Score 678; DB 10; Length 678;  
Best Local Similarity 100.0%; Pred. No. 9.8e-208;  
Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGGTCTTCCAAGATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGATGGAAGGA 60  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1 ATGAGGTCTTCCAAGATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGATGGAAGGA 60  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 61 ACGGTCAATGGGCACGAGTTTGAAATAGAGCGCAAGGAGAGGGGAGGCCATACGAAGGC 120  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 61 ACGGTCAATGGGCACGAGTTTGAAATAGAGCGCAAGGAGAGGGGAGGCCATACGAAGGC 120  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTGGGATATT 180  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTGGGATATT 180  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 181 TTGTCAACCAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA 240  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 181 TTGTCAACCAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA 240  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 241 GACTATAAAAAGCTGTCAATTTCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTGAA 300  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 241 GACTATAAAAAGCTGTCAATTTCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTGAA 300  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 301 GACGGTGGCGTCTGTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 301 GACGGTGGCGTCTGTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 361 AAGGTCAAGTTTCATTTGGCGTGAACCTTCTCCGATGGACCTGTTATGCAAAAGAAGACA 420  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 361 AAGGTCAAGTTTCATTTGGCGTGAACCTTCTCCGATGGACCTGTTATGCAAAAGAAGACA 420  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTTTGAAGGAGAG 480  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTTTGAAGGAGAG 480  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 481 ATTATAAGGCTCTGAAGCTGAAAGACGGTGGTTCATCTAGTTGAATTCAAAAGTATT 540  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 481 ATTATAAGGCTCTGAAGCTGAAAGACGGTGGTTCATCTAGTTGAATTCAAAAGTATT 540  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 541 TACATGGCAAAGAAGCCCTGTGCAGCTACCAAGGGTACTACTATGTTGACTCCAACTGGAT 600  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 541 TACATGGCAAAGAAGCCCTGTGCAGCTACCAAGGGTACTACTATGTTGACTCCAACTGGAT 600  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCAGGGGACGC 660  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCAGGGACGC 660  
QY 661 CACCATCTGTTCCCTTTAA 678  
DB ||||||||||||||||||||||||  
Db 661 CACCATCTGTTCCCTTTAA 678

RESULT 6

ADL46204  
ID ADL46204 standard; DNA; 678 BP.  
XX  
AC ADL46204;  
XX  
DT 20-MAY-2004 (first entry)  
XX  
DE Discosoma red fluorescent protein (DsRed) coding sequence.  
XX  
KW ds; gene; red fluorescent protein; DsRed; fluorescence; red wavelength;  
KW oligomerization; tetramerization; immunoassay; hybridization assay.  
XX  
OS Discosoma sp.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..678  
FT /\*tag= a  
FT /product= "DsRed protein"

PN WO2003086446-A1.  
XX  
PD 23-OCT-2003.  
XX  
PF 09-APR-2003; 2003WO-US010879.  
XX  
PR 10-APR-2002; 2002US-00121258.  
PR 29-JUL-2002; 2002US-00209208.  
XX  
PA (REGC ) UNIV CALIFORNIA.  
XX  
PI Tsien RY, Campbell RE, Baird GS;  
XX  
DR WPI; 2003-845265/78.  
DR P-PSDB; ADL46203.  
XX

PT New monomeric and dimeric Anthozoan fluorescent protein variants with  
PT reduced propensity to oligomerize, and encoding polynucleotides, useful  
PT in molecular biology, e.g. in immunoassays or in tracking protein  
PT movement in cells.

XX Disclosure; SEQ ID NO 2; 166pp; English.

XX The invention relates to a polynucleotide sequence encoding a Discosoma  
CC red fluorescent protein (DsRed) variant having a reduced propensity to  
CC oligomerize. The protein variant comprises one or more amino acid  
CC substitutions at the AB and/or AC interface(s) of the wild-type DsRed  
CC sequence, where the substitutions result in reduced propensity of the  
CC DsRed variant to form tetramers and where the variant displays detectable  
CC fluorescence of at least one red wavelength. The composition and methods  
CC are useful in producing red fluorescent proteins having reduced  
CC propensity for oligomerization, especially tetramerization. The protein  
CC may be used in molecular biology and in other scientific applications,  
CC such as in immunoassays or hybridization assays, or in tracking the  
CC movement of proteins in cells. This sequence corresponds to the DsRed  
CC coding sequence.

XX Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Query Match 100.0%; Score 678; DB 11; Length 678;  
Best Local Similarity 100.0%; Pred. No. 9.8e-208;  
Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGGTCTTCCAAGATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGATGGAAGGA 60  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 1 ATGAGGTCTTCCAAGATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGATGGAAGGA 60





Db 421 ATGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
Qy 481 ATTCTAAGGCTCTGAAGCTGAAAGACGGTGGTCAATACCTAGTTGAATTCAAAAGTATT 540  
Db 481 ATTCTAAGGCTCTGAAGCTGAAAGACGGTGGTCAATACCTAGTTGAATTCAAAAGTATT 540  
Qy 541 TACATGGCAAGAAGCCCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAT 600  
Db 541 TACATGGCAAGAAGCCCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAT 600  
Qy 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC 660  
Db 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC 660  
Qy 661 CACCATCTGTTCTTTAA 678  
Db 661 CACCATCTGTTCTTTAA 678

RESULT 8  
ADI36420  
ID ADI36420 standard; DNA; 678 BP.  
XX  
AC ADI36420;  
XX  
DT 22-APR-2004 (first entry)  
XX  
DE Discosoma sp. red fluorescent protein (RED) DNA.  
XX  
KW Fluorobody; binding ligand; green fluorescent protein; GFP;  
KW target detection; gene; red fluorescent protein; ds; RED.  
XX  
OS Discosoma sp.

Key Location/Qualifiers  
FH 1..678  
FT /\*tag= a  
FT /product= "Discosoma sp. red fluorescent protein (RED)"  
XX  
PN US2003203355-A1.  
XX  
PD 30-OCT-2003.  
XX  
PF 24-APR-2002; 2002US-00132067.  
XX  
PR 24-APR-2002; 2002US-00132067.  
XX  
PA (LALA-) LOS ALAMOS NAT LAB.  
XX (REGC ) UNIV CALIFORNIA.

PI Bradbury AM, Zeytun A, Waldo GS;  
XX  
XX WPI; 2004-154325/15.  
DR P-PSDB; ADI36421.  
XX  
PT Novel binding ligand with intrinsic fluorescence and comprising  
PT fluorescent protein having heterologous binding sites, useful for  
PT detecting target molecule.  
XX  
PS Example 6; SEQ ID NO 3; 23pp; English.  
XX  
CC The invention relates to binding ligands (fluorobodies) with intrinsic  
CC fluorescence, which comprises green fluorescent protein (GFP) having  
CC heterologous binding sites. The binding ligand is useful for detecting  
CC the target molecule and is efficiently detects the target molecule. The  
CC present sequence is Discosoma sp. red fluorescent protein (RED) DNA used  
CC in the exemplification of the invention.  
XX  
SQ Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Query Match 100.0%; Score 678; DB 12; Length 678;  
Best Local Similarity 100.0%; Pred. No. 9.8e-208;  
Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTTCATGAGGTTTAAAGTTCGCATGGAAGGA 60  
Db 1 ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTTCATGAGGTTTAAAGTTCGCATGGAAGGA 60  
Qy 61 ACGGTCAATGGCACGAGTTTGAATAGAAAGCGAAGGAGAGGGAGGCCATACGAAGGC 120  
Db 61 ACGGTCAATGGCACGAGTTTGAATAGAAAGCGAAGGAGAGGGAGGCCATACGAAGGC 120  
Qy 121 CACAATACCGTAAAGCTTAAGGTAAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
Db 121 CACAATACCGTAAAGCTTAAGGTAAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
Qy 181 TTGTCAACCAATTTTCAAGTATGGAAGCAAGGTATATGTCAAGCACCTTCCGACATACCA 240  
Db 181 TTGTCAACCAATTTTCAAGTATGGAAGCAAGGTATATGTCAAGCACCTTCCGACATACCA 240  
Qy 241 GACTATAAAAAGCTGTCTATTTCCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300  
Db 241 GACTATAAAAAGCTGTCTATTTCCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300  
Qy 301 GACGTGGCGTCTGTACTGTAAACCCAGGATTTCCAGTTCAGGATGCGTGTTCATCTAC 360  
Db 301 GACGTGGCGTCTGTACTGTAAACCCAGGATTTCCAGTTCAGGATGCGTGTTCATCTAC 360  
Qy 361 AAGGTCAAGTTCATTTGGCGTGAACCTTTCCCTCCGATGGACCTGTTATGCAAAAAGAGACA 420  
Db 361 AAGGTCAAGTTCATTTGGCGTGAACCTTTCCCTCCGATGGACCTGTTATGCAAAAAGAGACA 420  
Qy 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTATGGCGTGTGAAAGGAGAG 480  
Db 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTATGGCGTGTGAAAGGAGAG 480  
Qy 481 ATTCTAAGGCTCTGAAGCTGAAAGACGGTGGTCTATTACCTAGTTGAATTCAAAAGTATT 540  
Db 481 ATTCTAAGGCTCTGAAGCTGAAAGACGGTGGTCTATTACCTAGTTGAATTCAAAAGTATT 540  
Qy 541 TACATGGCAAGAAGCCCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAT 600  
Db 541 TACATGGCAAGAAGCCCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAT 600  
Qy 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC 660  
Db 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC 660  
Qy 661 CACCATCTGTTCTTTAA 678  
Db 661 CACCATCTGTTCTTTAA 678

RESULT 9  
ADM97768  
ID ADM97768 standard; DNA; 678 BP.  
XX  
AC ADM97768;  
XX  
DT 01-JUL-2004 (first entry)  
XX  
DE D sp red fluorescent protein coding sequence SEQ ID NO: 21.  
XX  
KW ds; gene; enzyme; sensor cell; fluorescent protein;  
KW signal transduction detection system; promoter; targeting sequence;  
XX targeted drug.  
OS Discosoma sp.  
XX  
FH Key Location/Qualifiers  
FT 1..678  
FT /\*tag= a  
FT /product= "fluorescent protein"  
XX  
PN WO2004031415-A2.  
XX



PD 15-APR-2004.  
XX  
PF 05-SEP-2003; 2003WO-US028078.  
XX  
PR 05-SEP-2002; 2002US-0408297P.  
XX (VERT-) VERTEX PHARM INC.  
PA  
PI Whitney MA, Zeh K, Sanders PS;  
XX  
XX WPI; 2004-330208/30.  
DR P-PSDB; ADM97769.  
XX  
XX Developing a sensor cell, useful in determining the activity of a target gene and in developing therapeutic drugs, comprises providing cells comprising a signal transduction detection system and introducing DNA construct into cells.  
XX  
PS Disclosure; Page 167-168; 234pp; English.  
XX  
CC The present invention relates to a method of developing a sensor cell, for determining the activity of a target gene in the cell, which comprises providing a homogeneous population of cells, where each of the cells comprises a signal transduction detection system and introducing into the population of cells an isolated DNA construct comprising a promoter operatively linked to a targeting sequence. The method is useful in developing a sensor cell for determining the activity of a target gene in the cell. The sensor cell and the methods are useful in developing new and therapeutic drugs directed to the targets. The present sequence is a coding sequence shown in the exemplification of the invention.  
XX  
SQ Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;  
  
Query Match 100.0%; Score 678; DB 12; Length 678;  
Best Local Similarity 100.0%; Pred. No. 9.8e-208;  
Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTCCGATGGAAGGA 60  
Db |  
1 ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTCCGATGGAAGGA 60  
  
QY 61 ACGGTCAATGGGCACGAGTTTGAAATAGAAGCGCAAGAGAGGGGAGGCCATACGAAGGC 120  
Db |  
61 ACGGTCAATGGGCACGAGTTTGAAATAGAAGCGCAAGAGAGGGGAGGCCATACGAAGGC 120  
  
QY 121 CACAATACCGTAAAGCTTAAAGGTAAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
Db |  
121 CACAATACCGTAAAGCTTAAAGGTAAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
  
QY 181 TTGTCACCAACAATTTTCAGTATGGAAGCAAGGTATATGTCGAAGCACCCCTGCCACATACCA 240  
Db |  
181 TTGTCACCAACAATTTTCAGTATGGAAGCAAGGTATATGTCGAAGCACCCCTGCCACATACCA 240  
  
QY 241 GACTATAAAAGCTGTCATTTCTCTGAAGGATTTTAATGGGAAAGGTCATGAACTTTGAA 300  
Db |  
241 GACTATAAAAGCTGTCATTTCTCTGAAGGATTTTAATGGGAAAGGTCATGAACTTTGAA 300  
  
QY 301 GACGGTGGCGTCGTTACGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360  
Db |  
301 GACGGTGGCGTCGTTACGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360  
  
QY 361 AAGGTCAAGTTTCATTGGCGTGAACCTTTCCTCCGATGGACCTGTTATGCAAAGAAGACA 420  
Db |  
361 AAGGTCAAGTTTCATTGGCGTGAACCTTTCCTCCGATGGACCTGTTATGCAAAGAAGACA 420  
  
QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGGTTTGATATCCTCGTGATGGCGTGTGTTAAAGGAGAG 480  
Db |  
421 ATGGGCTGGGAAGCCAGCACTGAGCGGTTTGATATCCTCGTGATGGCGTGTGTTAAAGGAGAG 480  
  
QY 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTTCATTACCTAGTTGAATTCAAAAGTATT 540  
Db |  
481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTTCATTACCTAGTTGAATTCAAAAGTATT 540

QY 541 TACATGGCAAAAGAGCCCTGTGCGAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600  
Db |  
541 TACATGGCAAAAGAGCCCTGTGCGAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600  
  
QY 601 ATAACAAGCCACAACGAAGACTATACAATCGTTTGAGCAGTATGAAGAACCAGGAGGACGC 660  
Db |  
601 ATAACAAGCCACAACGAAGACTATACAATCGTTTGAGCAGTATGAAGAACCAGGAGGACGC 660  
  
QY 661 CACCATCTGTTCTTTAA 678  
Db |  
661 CACCATCTGTTCTTTAA 678  
  
RESULT 10  
AAL47952  
ID AAL47952 standard; DNA; 859 BP.  
XX  
AC AAL47952;  
XX  
DT 26-SEP-2002 (first entry)  
XX  
DE Discosoma red fluorescent protein coding sequence.  
XX  
KW Yeast; RAD54; promoter; genotoxicity cassette; cytotoxicity cassette; modified yeast strain; environmental pollution; gene; ds.  
KW  
XX Discosoma sp.  
XX  
FH Key Location/Qualifiers  
FT CDS 54..731  
FT /\*tag= a  
FT /product= "red fluorescent protein"  
XX  
PN DE10061872-A1.  
XX  
PD 20-JUN-2002.  
XX  
PF 12-DEC-2000; 2000DE-01061872.  
XX  
PR 12-DEC-2000; 2000DE-01061872.  
XX  
PA (LICH/) LICHTENBERG-FRATE H.  
XX  
PI Lichtenberg-Frate H;  
XX  
DR WPI; 2002-539633/58.  
DR P-PSDB; AAO18270.  
XX  
PT Modified yeast strain, useful for detecting toxic compounds in environment, contains integrated cassettes responsive to genotoxic and cytotoxic compounds.  
PT  
XX  
PS Disclosure; Page 20-21; 34pp; German.  
XX  
CC The present invention relates to a modified yeast strain that contains, integrated stably and functionally in its genome, a genotoxicity cassette and a cytotoxicity cassette, each comprising a promoter and reporter gene, both of which are different in the two cassettes. The modified yeast strain is used to detect environmental pollution, especially genotoxic and/or cytotoxic substances in complex environmental contaminants, especially organic compounds, but also (non-)ionising radiation and chemical carcinogens. Particular applications are in monitoring (waste) water (e.g. as an early warning system), medical toxicology screening and for industrial process control. The present sequence is a marker gene suitable for use in the cassettes of the present invention  
CC  
XX  
SQ Sequence 859 BP; 263 A; 163 C; 228 G; 205 T; 0 U; 0 Other;  
  
Query Match 100.0%; Score 678; DB 6; Length 859;  
Best Local Similarity 100.0%; Pred. No. 1.1e-207;  
Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTTCATGAGTTTAAGGTTTCGCATGGAAGGA 60  
Db |||||||  
QY 54 ATGAGGTCTTCCNAGAATGTTATCAAGGAGTTTCATGAGTTTAAGGTTTCGCATGGAAGGA 113  
Db |||||||  
QY 61 ACGTCAATGGGCACGAGTTTGAAATAGAACGCGAAGGAGAGGGAGGCCATACGAAGGC 120  
Db |||||||  
QY 114 ACGTCAATGGGCACGAGTTTGAAATAGAACGCGAAGGAGAGGGAGGCCATACGAAGGC 173  
QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
Db |||||||  
QY 174 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGGACCTTTGCCATTTGCTTGGGATATT 233  
QY 181 TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240  
Db |||||||  
QY 234 TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 293  
QY 241 GACTATAAAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAAGGTCATGAATTTGAA 300  
Db |||||||  
QY 294 GACTATAAAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAAGGTCATGAATTTGAA 353  
QY 301 GACGGTGGCGTCGTACTCTGAACCCAGGATTCAGTTTGAGGATGGCTGTTTCATCTAC 360  
Db |||||||  
QY 354 GACGGTGGCGTCGTACTCTGAACCCAGGATTCAGTTTGAGGATGGCTGTTTCATCTAC 413  
QY 361 AAGGTCAAGTTTCATTTGGCGTGAACCTTTCCTTCCGATGGACCTGTTATGCAAAAGAAGACA 420  
Db |||||||  
QY 414 AAGGTCAAGTTTCATTTGGCGTGAACCTTTCCTTCCGATGGACCTGTTATGCAAAAGAAGACA 473  
QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
Db |||||||  
QY 474 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 533  
QY 481 ATTCATAAGGCTCTGAAGCTGAAAGACGCGTGGTCAATACCTAGTTGAATTCAAAAGTATT 540  
Db |||||||  
QY 534 ATTCATAAGGCTCTGAAGCTGAAAGACGCGTGGTCAATACCTAGTTGAATTCAAAAGTATT 593  
QY 541 TACATGGCAAAAGACCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600  
Db |||||||  
QY 594 TACATGGCAAAAGACCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 653  
QY 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCCGAGGGACGC 660  
Db |||||||  
QY 654 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCCGAGGGACGC 713  
QY 661 CACCATCTGTTCCCTTTAA 678  
Db |||||||  
QY 714 CACCATCTGTTCCCTTTAA 731

RESULT 11

AAD53432  
ID AAD53432 standard; DNA; 859 BP.  
XX  
AC AAD53432;  
XX  
DT 28-MAY-2003 (first entry)  
XX  
DE Discosoma species red fluorescent protein (RFP) encoding DNA.  
XX  
KW phosphorylation indicator; fluorescent protein; detection; phosphatase;  
KW kinase; red fluorescent protein; RFP; gene; ds.  
XX  
OS Discosoma sp.  
XX  
FH Key Location/Qualifiers  
FT CDS 54..731  
FT /\*tag= a  
FT /product= "Discosoma sp. red fluorescent protein (RFP)"  
XX  
PN WO200295058-A2.

28-NOV-2002.

XX

PF 24-MAY-2002; 2002WO-US016955.  
XX  
PR 24-MAY-2001; 2001US-00865291.  
XX  
PA (REGC ) UNIV CALIFORNIA.  
XX  
PI Tsien RY, Ting AY, Zhang J;  
XX  
DR WPI; 2003-148474/14.  
DR P-PSDB; AAE34962.  
XX  
PT Novel chimeric phosphorylation indicators, useful for detecting  
PT kinase/phosphatase in samples, has donor molecule, phosphorylatable  
PT domain, phosphoaminoacid binding domain, and acceptor molecule, in  
PT operative linkage.  
XX  
PS Disclosure; Col 64-65; 38pp; English.  
XX  
CC The present invention relates to chimeric phosphorylation indicators  
CC comprising a phosphorylation polypeptide and a fluorescent protein or in  
CC operative linkage, a donor molecule, a phosphorylatable domain, a  
CC phosphoaminoacid binding domain (PAABD) and an acceptor molecule. The  
CC phosphorylation indicators of the invention are useful for detecting  
CC kinases or phosphatases in a biological sample. They are also useful in  
CC high throughput analysis e.g. for detecting a kinase inhibitor or  
CC phosphatase inhibitor. The present sequence is Discosoma species red  
CC fluorescent protein (RFP) encoding DNA used in the invention  
XX  
SQ Sequence 859 BP; 263 A; 163 C; 228 G; 205 T; 0 U; 0 Other;

Query Match 100.0%; Score 678; DB 8; Length 859;  
Best Local Similarity 100.0%; Pred. NO. 1.1e-207;  
Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTTCATGAGTTTAAGGTTTCGCATGGAAGGA 60  
Db |||||||  
QY 54 ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTTCATGAGTTTAAGGTTTCGCATGGAAGGA 113  
Db |||||||  
QY 61 ACGTCAATGGGCACGAGTTTGAAATAGAACGCGAAGGAGAGGGAGGCCATACGAAGGC 120  
Db |||||||  
QY 114 ACGTCAATGGGCACGAGTTTGAAATAGAACGCGAAGGAGAGGGAGGCCATACGAAGGC 173  
QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
Db |||||||  
QY 174 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGGACCTTTGCCATTTGCTTGGGATATT 233  
QY 181 TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240  
Db |||||||  
QY 234 TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 293  
QY 241 GACTATAAAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAAGGTCATGAATTTGAA 300  
Db |||||||  
QY 294 GACTATAAAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAAGGTCATGAATTTGAA 353  
QY 301 GACGGTGGCGTCGTACTCTGAACCCAGGATTCAGTTTGAGGATGGCTGTTTCATCTAC 360  
Db |||||||  
QY 354 GACGGTGGCGTCGTACTCTGAACCCAGGATTCAGTTTGAGGATGGCTGTTTCATCTAC 413  
QY 361 AAGGTCAAGTTTCATTTGGCGTGAACCTTTCCTTCCGATGGACCTGTTATGCAAAAGAAGACA 420  
Db |||||||  
QY 414 AAGGTCAAGTTTCATTTGGCGTGAACCTTTCCTTCCGATGGACCTGTTATGCAAAAGAAGACA 473  
QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
Db |||||||  
QY 474 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 533  
QY 481 ATTCATAAGGCTCTGAAGCTGAAAGACGCGTGGTCAATACCTAGTTGAATTCAAAAGTATT 540  
Db |||||||  
QY 534 ATTCATAAGGCTCTGAAGCTGAAAGACGCGTGGTCAATACCTAGTTGAATTCAAAAGTATT 593  
QY 541 TACATGGCAAAAGACCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600  
Db |||||||  
QY 594 TACATGGCAAAAGACCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 653

QY 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGACGAGTATGAAAGAACCGAGGGACGC 660  
| | | | |  
Db 654 ATAACAAGCCACAACGAAGACTATACAATCGTTGACGAGTATGAAAGAACCGAGGGACGC 713  
| | | | |  
QY 661 CACCATCTGTTCCCTTTAA 678  
| | | | |  
Db 714 CACCATCTGTTCCCTTTAA 731  
| | | | |  
RESULT 12  
AAD61969  
ID AAD61969 standard; cDNA; 859 BP.  
XX  
AC AAD61969;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE Discosoma sp. red fluorescent protein (RFP) cDNA.  
XX  
KW Fluorescent protein; resonance energy transfer; pH; detection;  
KW red fluorescent protein; RFP; gene; ss.  
XX  
OS Discosoma sp.  
XX  
FH Key Location/Qualifiers  
FT CDS 54..731  
FT /\*tag= a  
FT /product= "Discosoma sp. red fluorescent protein"  
XX  
PN US2003170911-A1.  
XX  
PD 11-SEP-2003.  
XX  
PF 26-FEB-2001; 2001US-00794308.  
XX  
PR 26-FEB-2001; 2001US-00794308.  
XX  
PA (TSIE/) TSIENT R Y.  
PA (ZACH/) ZACHARIAS D A.  
PA (BAIR/) BAIRD G S.  
XX  
PI Tsien RY, Zacharias DA, Baird GS;  
XX  
DR WPI; 2003-802418/75.  
DR P-PSDB; ABW00918.  
XX  
PT Fluorescent proteins containing a mutation that reduces or eliminates its  
PT ability to oligomerize which gives more reliable fluorescence resonance  
PT energy transfer results and are useful to detect molecule interaction,  
PT enzymes, or sample pH.  
XX  
PS Disclosure; Page 29-30; Opp; English.  
XX  
CC The invention relates to a non-oligomerising fluorescent protein  
CC containing a mutation that reduces or eliminates its ability to  
CC oligomerise. The fluorescent protein gives more reliable fluorescence  
CC resonance energy transfer results and are useful to detect molecule  
CC interaction, enzymes, or sample pH. These are also used to identify  
CC agents or conditions that regulate expression of control sequences. The  
CC present sequence is Discosoma sp. red fluorescent protein (RFP) cDNA  
XX  
SQ Sequence 859 BP; 263 A; 163 C; 228 G; 205 T; 0 U; 0 Other;  
  
Query Match 100.0%; Score 678; DB 10; Length 859;  
Best Local Similarity 100.0%; Pred. No. 1.1e-207;  
Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGATGGAAGGA 60  
| | | | |  
Db 54 ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGATGGAAGGA 113  
| | | | |  
QY 61 ACGGTCAATGGGCACGAGTTTGAATAAGGCGGAAGGAGAGGGGAGGCCATACGAAGGC 120  
| | | | |

Db 114 ACGGTCAATGGGCACGAGTTTGAATAAGGCGGAAGGAGAGGGGAGGCCATACGAAGGC 173  
| | | | |  
QY 121 CACAATACCGTAAAGCTTAAGGTAAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
| | | | |  
Db 174 CACAATACCGTAAAGCTTAAGGTAAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 233  
| | | | |  
QY 181 TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240  
| | | | |  
Db 234 TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 293  
| | | | |  
QY 241 GACTATAAAAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAA 300  
| | | | |  
Db 294 GACTATAAAAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAA 353  
| | | | |  
QY 301 GACGGTGGCGTCGTTACTGTAAACCCAGGATTCCAGTTTGCAGGATGGCTGTTTCATCTAC 360  
| | | | |  
Db 354 GACGGTGGCGTCGTTACTGTAAACCCAGGATTCCAGTTTGCAGGATGGCTGTTTCATCTAC 413  
| | | | |  
QY 361 AAGGTCAGGTTCAATGGCGTGAACTTCCCTCCGATGGACCTGTATGCAAAAAGAGACA 420  
| | | | |  
Db 414 AAGGTCAGGTTCAATGGCGTGAACTTCCCTCCGATGGACCTGTATGCAAAAAGAGACA 473  
| | | | |  
QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
| | | | |  
Db 474 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 533  
| | | | |  
QY 481 ATTCAATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAAGTATT 540  
| | | | |  
Db 534 ATTCAATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAAGTATT 593  
| | | | |  
QY 541 TACATGGCAAAAGAACCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600  
| | | | |  
Db 594 TACATGGCAAAAGAACCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 653  
| | | | |  
QY 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC 660  
| | | | |  
Db 654 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC 713  
| | | | |  
QY 661 CACCATCTGTTCCCTTTAA 678  
| | | | |  
Db 714 CACCATCTGTTCCCTTTAA 731  
| | | | |

RESULT 13  
ADF70404  
ID ADF70404 standard; cDNA; 678 BP.  
XX  
AC ADF70404;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE Discosoma wild-type GFP variant cDNA SeqID27.  
XX  
KW ligand; orphan receptor protein; fusion protein; fluorescent protein;  
KW cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;  
KW GFPuv; Enhanced GFP; EGFP; gene; ss.  
XX  
OS Discosoma sp.  
XX  
PN WO2003071272-A1.  
XX  
PD 28-AUG-2003.  
XX

PF 21-FEB-2003; 2003WO-JP001901.  
XX  
PR 22-FEB-2002; 2002JP-00045728.  
PR 23-JUL-2002; 2002JP-00213949.  
PR 11-OCT-2002; 2002JP-00298237.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Hinuma S, Fujii R, Ogi K, Komatsu H, Kawamata Y, Hosoya M;



XX WPI; 2003-697654/66.  
DR P-PSDB; ADF70403.  
XX  
PT Transformation of cells with a fusion protein of an orphan receptor  
PT protein with a fluorescent protein useful for identification of ligands  
PT to the orphan receptor.  
XX  
PS Disclosure; SEQ ID NO 27; 594pp; Japanese.  
XX  
CC This invention relates to a novel method of identifying ligands to an  
CC orphan receptor protein which comprises transforming cells with DNA  
CC encoding a fusion protein of the orphan receptor with a fluorescent  
CC protein, so that the fusion protein is expressed in the cells (or cell  
CC membranes isolated from them) and contacting the cells with the potential  
CC ligand to be tested. A suitable fluorescent protein (GFP), for example GFP-1,  
CC the fusion protein is green fluorescent protein (GFP), for example GFP-1,  
CC wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the  
CC identification of ligands binding to an orphan receptor protein.  
XX  
SQ Sequence 678 BP; 204 A; 129 C; 179 G; 166 T; 0 U; 0 Other;  
  
Query Match 99.9%; Score 677; DB 10; Length 678;  
Best Local Similarity 100.0%; Pred. No. 2.1e-207;  
Matches 677; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 ATGAGGTCTTCCAAGATGTTATCAAGGAGTTTCATGAGTTTAAGGTTTCGATGGAAGGA 60  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1 ATGAGGTCTTCCAAGATGTTATCAAGGAGTTTCATGAGTTTAAGGTTTCGATGGAAGGA 60  
  
QY 61 ACGGTCAATGGGCACGAGTTTGAATAGAAAGCGAGGAGAGGGGAGGCCATACGAAGGC 120  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
61 ACGGTCAATGGGCACGAGTTTGAATAGAAAGCGAGGAGAGGGGAGGCCATACGAAGGC 120  
  
QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
  
QY 181 TTGTCAACCAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA 240  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
181 TTGTCAACCAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA 240  
  
QY 241 GACTATAAAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
241 GACTATAAAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300  
  
QY 301 GACGGTGGCGTCTACTGTAAACCCAGGATTCAGATTCCAGGATGGCTGTTTCATCTAC 360  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
301 GACGGTGGCGTCTACTGTAAACCCAGGATTCAGATTCCAGGATGGCTGTTTCATCTAC 360  
  
QY 361 AAGGTCAAGTTTCATGGCGTGAACCTTTCCTCCGATGGACCTGTTATGCAAAAGAGACA 420  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
361 AAGGTCAAGTTTCATGGCGTGAACCTTTCCTCCGATGGACCTGTTATGCAAAAGAGACA 420  
  
QY 421 ATGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
421 ATGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
  
QY 481 ATTCAATAGGCTCTGAAGCTGAAAGACGGTGGTCACTTACCTAGTTGAATTCAAAAGTATT 540  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
481 ATTCAATAGGCTCTGAAGCTGAAAGACGGTGGTCACTTACCTAGTTGAATTCAAAAGTATT 540  
  
QY 541 TACATGGCAAAGAGCCTGTGCAGCTACCCAGGGTACTACTATGTTGACTCCAAACTGGAT 600  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
541 TACATGGCAAAGAGCCTGTGCAGCTACCCAGGGTACTACTATGTTGACTCCAAACTGGAT 600  
  
QY 601 ATAACAAGCCACAACGAAGACTATACAAATCGTTGACGAGTATGAAAGAACCGAGGGACGC 660  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
601 ATAACAAGCCACAACGAAGACTATACAAATCGTTGACGAGTATGAAAGAACCGAGGGACGC 660  
  
QY 661 CACCATCTGTTCCCTTTA 677  
Db ||||||||||||||||||  
661 CACCATCTGTTCCCTTTA 677

RESULT 14  
ACA62995  
ID ACA62995 standard; DNA; 3311 BP.  
XX  
AC ACA62995;  
XX  
DT 23-SEP-2003 (first entry)  
XX  
DE Plasmid DNA containing coding sequence for RFP.  
XX  
KW Fluorescently-tagged enzyme; substrate; cell population; GFP;  
KW quantification of enzymatic activity; green fluorescent protein;  
KW red fluorescent protein; RFP; enhanced green fluorescent protein; EGFP;  
KW enzymatic process; cellular enzyme; chemotherapeutic drug;  
KW multidrug resistance; MDR; ds.  
XX  
OS Aequorea victoria.  
OS Synthetic.  
XX  
PN US2003049597-A1.  
XX  
PD 13-MAR-2003.  
XX  
PF 01-MAR-2001; 2001US-00797496.  
XX  
PR 01-MAR-2001; 2001US-00797496.  
XX  
PA (SIMO/) SIMON S M.  
PA (CHEN/) CHEN Y.  
XX  
PI Simon SM, Chen Y;  
XX  
WPI; 2003-555145/52.  
XX  
PT Simultaneously quantifying in situ the relationship between an enzyme and  
PT its substrate for study of enzymatic processes at a cellular level,  
PT comprises optical measurements on cells expressing the fluorescently  
PT tagged enzyme.  
XX  
PS Disclosure; Page 8-9; 21pp; English.  
XX  
CC The present invention relates to a method for simultaneously quantifying  
CC in situ the relationship between a fluorescently-tagged enzyme (E) and  
CC its substrate in a population of cells. A population of cells expressing  
CC (E) is created in which different cells contain different amounts of (E).  
CC The cells are incubated with a substrate, and the concentration and  
CC enzymatic activity of (E) in each cell are simultaneously quantified by  
CC optical means. The fluorescently-tagged enzyme is produced by linking the  
CC enzyme with Aequorea victoria green fluorescent protein (GFP), or its  
CC variants or derivatives such as red fluorescent protein (RFP) and  
CC enhanced green fluorescent protein (EGFP). The method of the invention is  
CC useful for simultaneously quantifying in situ the relationship between an  
CC enzyme and its substrate in a cell. The method is useful for the study of  
CC enzymatic processes at the cellular level, and especially for examining  
CC the interactions of cellular enzymes with chemotherapeutic drugs, e.g.  
CC for elucidating mechanisms of multidrug resistance (MDR). The present  
CC sequence represents a plasmid containing the coding sequence for RFP  
XX  
SQ Sequence 3311 BP; 858 A; 778 C; 860 G; 815 T; 0 U; 0 Other;  
  
Query Match 99.9%; Score 677; DB 9; Length 3311;  
Best Local Similarity 100.0%; Pred. No. 4.6e-207;  
Matches 677; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 ATGAGGTCTTCCAAGATGTTATCAAGGAGTTTCATGAGTTTAAAGTTTCGATGGAAGGA 60  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
289 ATGAGGTCTTCCAAGATGTTATCAAGGAGTTTCATGAGTTTAAAGTTTCGATGGAAGGA 348  
  
QY 61 ACGGTCAATGGGCACGAGTTTGAATAGAAAGCGAGGAGGGGAGGCCATACGAAGGC 120  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
349 ACGGTCAATGGGCACGAGTTTGAATAGAAAGCGAGGAGGGGAGGCCATACGAAGGC 408



QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
|||||  
Db 409 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 468  
|||||  
QY 181 TTGTCAACCACAATTTCAAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA 240  
|||||  
Db 469 TTGTCAACCACAATTTCAAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA 528  
|||||  
QY 241 GACTATAAAAAAGCTGTAATTCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300  
|||||  
Db 529 GACTATAAAAAAGCTGTAATTCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 588  
|||||  
QY 301 GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360  
|||||  
Db 589 GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 648  
|||||  
QY 361 AAGGTCAAGTTCATTGGCGTGAACCTTTCCTTCCGATGGACCTGTTATGCAAAAGAAAGACA 420  
|||||  
Db 649 AAGGTCAAGTTCATTGGCGTGAACCTTTCCTTCCGATGGACCTGTTATGCAAAAGAAAGACA 708  
|||||  
QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTCCTGATCCCTCGTGATGGCGTGTGAAAGGAGAG 480  
|||||  
Db 709 ATGGGCTGGGAAGCCAGCACTGAGCGTTCCTGATCCCTCGTGATGGCGTGTGAAAGGAGAG 768  
|||||  
QY 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT 540  
|||||  
Db 769 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT 828  
|||||  
QY 541 TACATGGCAAGAAGCCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600  
|||||  
Db 829 TACATGGCAAGAAGCCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 888  
|||||  
QY 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGGGGACGC 660  
|||||  
Db 889 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGGGGACGC 948  
|||||  
QY 661 CACCATCTGTTCCCTTTA 677  
|||||  
Db 949 CACCATCTGTTCCCTTTA 965  
|||||

RESULT 15  
ABL61142  
ID ABL61142 standard; DNA; 666 BP.  
XX  
AC ABL61142;  
XX  
DT 18-SEP-2002 (first entry)  
XX  
DE Red fluorescent protein reporter DNA fragment.  
XX  
KW Alpha-SMA; smooth muscle actin; screening; smooth muscle cell;  
XX myofibroblast; gene therapy; red fluorescent protein; ds.  
OS Synthetic.  
XX  
PN EP1172375-A1.  
XX  
PD 16-JAN-2002.  
XX  
PF 22-DEC-2000; 2000EP-00128446.  
XX  
PR 11-JUL-2000; 2000DE-01033633.  
PR 31-OCT-2000; 2000DE-01053879.  
XX  
PA (ODEN/) ODENTHAL M.  
XX  
PI Odenthal M, Jung D;  
XX  
DR WPI; 2002-149590/20.  
XX  
PT New nucleic acid containing regulatory region of the smooth muscle actin

PT gene, useful e.g. for manipulating gene expression in smooth muscle  
PT cells.  
XX  
PS Disclosure; Page 20; 44pp; German.  
XX  
CC This invention describes a novel nucleic acid (I) comprising: (i) at  
CC least one functional region (Ia) from the regulatory region of the alpha-  
CC smooth muscle actin (SMA) gene and (ii) at least one additional  
CC functional sequence (Ib) operably linked to (Ia). The products of the  
CC invention can be used for preparing genetically modified eukaryotic cells  
CC or organisms, for isolation and screening of smooth muscle cells,  
CC myofibroblasts or related cells, and for manipulation of gene expression  
CC and/or cell function in smooth muscle cell or myofibroblasts,  
CC particularly for gene therapy. Component (Ia) provides cell-type- or  
CC differentiation-specific expression or modulation of genes. This sequence  
CC represents a DNA fragment of red fluorescent protein (EGFP) which can be  
CC used as a reporter molecule under the control of the alpha-SMA described  
CC in the invention  
XX

SQ Sequence 666 BP; 202 A; 126 C; 176 G; 162 T; 0 U; 0 Other;

Query Match 98.1%; Score 665; DB 6; Length 666;  
Best Local Similarity 100.0%; Pred. No. 1.5e-203;  
Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 AAGAATGTTATCAAGGAGTTTCATGAGTTTAAAGTTCGCATGGAAGAACGGTCAATGGG 72  
|||||  
Db 1 AAGAATGTTATCAAGGAGTTTCATGAGTTTAAAGTTCGCATGGAAGAACGGTCAATGGG 60  
|||||  
QY 73 CACGAGTTTGAAAATAGAAAGCGAAGGAGAGGGGAGGCCATACGAAGGCCACAATACCGTA 132  
|||||  
Db 61 CACGAGTTTGAAAATAGAAAGCGAAGGAGAGGGGAGGCCATACGAAGGCCACAATACCGTA 120  
|||||  
QY 133 AAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATTTTGTCAACCACAA 192  
|||||  
Db 121 AAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATTTTGTCAACCACAA 180  
|||||  
QY 193 TTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCAGACTATAAAAG 252  
|||||  
Db 181 TTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCAGACTATAAAAG 240  
|||||  
QY 253 CTGTCAATTCCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAAGACGGTGGCGTC 312  
|||||  
Db 241 CTGTCAATTCCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAAGACGGTGGCGTC 300  
|||||  
QY 313 GTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTACAAGGTCAGTTC 372  
|||||  
Db 301 GTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTACAAGGTCAGTTC 360  
|||||  
QY 373 ATTGGCGTGAACCTTTCCTCCGATGGACCTGTTATGCAAAAAGAGACAATGGGCTGGGAA 432  
|||||  
Db 361 ATTGGCGTGAACCTTTCCTCCGATGGACCTGTTATGCAAAAAGAGACAATGGGCTGGGAA 420  
|||||  
QY 433 GCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAAGGAGAGATTTCATAAGGCT 492  
|||||  
Db 421 GCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAAGGAGAGATTTCATAAGGCT 480  
|||||  
QY 493 CTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATTACATGGCAAAG 552  
|||||  
Db 481 CTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATTACATGGCAAAG 540  
|||||  
QY 553 AAGCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGATATAACAAGCCAC 612  
|||||  
Db 541 AAGCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGATATAACAAGCCAC 600  
|||||  
QY 613 AACGAAGACTATACAATCGTTGAGCAGTATGAAAAGAACCGAGGGACGCCACCATCTGTTTC 672  
|||||  
Db 601 AACGAAGACTATACAATCGTTGAGCAGTATGAAAAGAACCGAGGGACGCCACCATCTGTTTC 660  
|||||  
QY 673 CTTTA 677  
|||||  
Db 661 CTTTA 665  
|||||

Search completed: July 1, 2005, 16:33:30  
Job time : 1033 secs

---

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2005, 16:16:18 ; Search time 571 Seconds  
(without alignments)  
7447.532 Million cell updates/sec

Title: US-10-006-922A-11  
Perfect score: 678  
Sequence: 1 atgaggtcttccaagaatgt.....gccaccatctgttcctttaa 678

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 6313374 seqs, 3136092125 residues

Total number of hits satisfying chosen parameters: 12626748

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA: \*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq: \*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq: \*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq: \*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq: \*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq: \*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq: \*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq: \*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq: \*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq: \*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq: \*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq: \*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq: \*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq: \*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq: \*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq: \*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq: \*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq: \*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq: \*  
19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq: \*  
20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq: \*  
21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq: \*  
22: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq: \*  
23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq: \*  
24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq: \*  
25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq: \*  
26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	678	100.0	678	9	US-09-967-772-6
2	678	100.0	678	13	US-10-006-922-11
3	678	100.0	678	14	US-10-121-258-2
4	678	100.0	678	17	US-10-132-067-3
5	678	100.0	678	17	US-10-335-517-6
6	678	100.0	678	17	US-10-334-288-6
7	678	100.0	678	19	US-10-311-030-5
					Sequence 6, Appli
					Sequence 11, Appl
					Sequence 2, Appli
					Sequence 3, Appli
					Sequence 6, Appli
					Sequence 6, Appli
					Sequence 5, Appli

8	678	100.0	678	21	US-10-656-029-21	Sequence 21, Appl
9	678	100.0	859	9	US-09-999-745-66	Sequence 66, Appl
10	678	100.0	859	10	US-09-866-538-11	Sequence 11, Appl
11	678	100.0	859	10	US-09-794-308-11	Sequence 11, Appl
12	678	100.0	859	10	US-09-865-291-11	Sequence 11, Appl
13	678	100.0	859	19	US-10-433-640-12	Sequence 12, Appl
14	678	100.0	859	20	US-10-885-988-11	Sequence 11, Appl
15	678	100.0	859	21	US-10-857-622-11	Sequence 11, Appl
16	677	99.9	678	21	US-10-505-486-27	Sequence 27, Appl
17	677	99.9	3311	10	US-09-797-496B-3	Sequence 3, Appli
18	665	98.1	666	19	US-10-332-733-22	Sequence 22, Appl
19	661	97.5	711	19	US-10-314-936-1	Sequence 1, Appli
20	661	97.5	711	19	US-10-314-936-3	Sequence 3, Appli
21	661	97.5	711	23	US-11-021-014-1	Sequence 1, Appli
22	661	97.5	711	23	US-11-021-014-3	Sequence 3, Appli
23	658.8	97.2	898	13	US-10-006-922-45	Sequence 45, Appl
24	642.8	94.8	678	18	US-10-423-688A-40	Sequence 40, Appl
25	566.4	83.5	876	13	US-10-006-922-17	Sequence 17, Appl
26	566.4	83.5	876	15	US-10-161-403-39	Sequence 39, Appl
27	450.6	66.5	723	14	US-10-152-296-1	Sequence 1, Appli
28	450.6	66.5	723	19	US-10-739-656-1	Sequence 1, Appli
29	449.8	66.3	681	19	US-10-311-030-8	Sequence 8, Appli
30	449.8	66.3	713	19	US-10-311-030-11	Sequence 11, Appl
31	449.8	66.3	713	19	US-10-311-030-12	Sequence 12, Appl
32	449.2	66.3	678	14	US-10-081-864-7	Sequence 7, Appli
33	449.2	66.3	678	16	US-10-315-920-1	Sequence 1, Appli
34	447.2	66.0	681	14	US-10-121-258-3	Sequence 3, Appli
35	447.2	66.0	681	14	US-10-121-258-23	Sequence 23, Appl
36	447.2	66.0	4692	15	US-10-161-403-29	Sequence 29, Appl
37	447.2	66.0	4692	19	US-10-433-640-16	Sequence 16, Appl
38	447.2	66.0	6984	13	US-10-001-189-45	Sequence 45, Appl
39	447.2	66.0	7910	21	US-10-169-050-20	Sequence 20, Appl
40	447.2	66.0	9320	19	US-10-471-065-20	Sequence 20, Appl
41	447.2	66.0	9658	19	US-10-609-019-4	Sequence 4, Appli
42	447.2	66.0	9678	19	US-10-609-019-3	Sequence 3, Appli
43	447.2	66.0	10263	19	US-10-609-019-2	Sequence 2, Appli
44	446	65.8	678	16	US-10-315-920-3	Sequence 3, Appli
45	445.8	65.8	1638	15	US-10-214-932-51	Sequence 51, Appl

ALIGNMENTS

RESULT 1  
US-09-967-772-6  
; Sequence 6, Application US/09967772  
; Patent No. US20020164577A1  
; GENERAL INFORMATION:  
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: TSJEN, Roger  
; APPLICANT: GONZALEZ, Jesus  
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY OPTICAL METHODS  
; FILE REFERENCE: REGEN1290-5  
; CURRENT APPLICATION NUMBER: US/09/967,772  
; CURRENT FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: US 09/459,956  
; PRIOR FILING DATE: 1999-12-13  
; PRIOR APPLICATION NUMBER: US 08/765,860  
; PRIOR FILING DATE: 1996-12-19  
; PRIOR APPLICATION NUMBER: PCT/ US96/09652  
; PRIOR FILING DATE: 1996-06-06  
; PRIOR APPLICATION NUMBER: US 08/481,977  
; PRIOR FILING DATE: 1995-06-07  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 6  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Discosoma sp "red"  
US-09-967-772-6  
Query Match 100.0%; Score 678; DB 9; Length 678;  
Best Local Similarity 100.0%; Pred. No. 5.9e-216;



Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGGTCTTCCAAAGATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGATGGAAGGA 60  
|||||  
Db 1 ATGAGGTCTTCCAAAGATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGATGGAAGGA 60

QY 61 ACGGTCAATGGCACGAGTTTGAAATAGAAAGGCGAAGGAGAGGGGAGGCATACGAAGGC 120  
|||||  
Db 61 ACGGTCAATGGCACGAGTTTGAAATAGAAAGGCGAAGGAGAGGGGAGGCATACGAAGGC 120

QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTTGCCATTTTGGGATATT 180  
|||||  
Db 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTTGCCATTTTGGGATATT 180

QY 181 TTGTCAACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240  
|||||  
Db 181 TTGTCAACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240

QY 241 GACTATAAAAAGCTGTCATTTCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300  
|||||  
Db 241 GACTATAAAAAGCTGTCATTTCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300

QY 301 GACGGTGGCGTCGTACTGTAAACCCAGGATTCAGTTTGAGGATGGCTTTCATCTAC 360  
|||||  
Db 301 GACGGTGGCGTCGTACTGTAAACCCAGGATTCAGTTTGAGGATGGCTTTCATCTAC 360

QY 361 AAGGTCAAGTTTCAATGGCGTGAACCTTTCCCTCCGATGGACCTGTTATGCAAAAAGAGACA 420  
|||||  
Db 361 AAGGTCAAGTTTCAATGGCGTGAACCTTTCCCTCCGATGGACCTGTTATGCAAAAAGAGACA 420

QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
|||||  
Db 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGATCCTCGTGATGGCGTGTGAAAGGAGAG 480

QY 481 ATTCTAAGGCTCTGAAGCTGAAAGACGGTGGTCATTAACCTAGTTGAAATCAAAAGTATT 540  
|||||  
Db 481 ATTCTAAGGCTCTGAAGCTGAAAGACGGTGGTCATTAACCTAGTTGAAATCAAAAGTATT 540

QY 541 TACATGGCAAAAGACCTGTGCAGCTACAGGGTACTACTATGTTGACTCCAAACTGGAT 600  
|||||  
Db 541 TACATGGCAAAAGACCTGTGCAGCTACAGGGTACTACTATGTTGACTCCAAACTGGAT 600

QY 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC 660  
|||||  
Db 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC 660

QY 661 CACCATCTGTTCCCTTTAA 678  
|||||  
Db 661 CACCATCTGTTCCCTTTAA 678

RESULT 2

US-10-006-922-11

; Sequence 11, Application US/10006922

; Publication No. US20020197676A1

; GENERAL INFORMATION:

; APPLICANT: Lukyanov, Sergey A

; APPLICANT: Fradkov, Arcady F.

; APPLICANT: Labas, Yulii A.

; APPLICANT: Matz, Mikhail V.

; APPLICANT: Tersikh, Alexey

; TITLE OF INVENTION: No. US20020197676A1el Chromophores/Fluorophores and

; FILE REFERENCE: CLON-035CIP

; CURRENT APPLICATION NUMBER: US/10/006,922

; CURRENT FILING DATE: 2001-12-04

; PRIOR APPLICATION NUMBER: 09/120,330

; PRIOR FILING DATE: 1998-12-11

; PRIOR APPLICATION NUMBER: 09/457,898

; PRIOR FILING DATE: 1999-12-09

; PRIOR APPLICATION NUMBER: 09/458,144

; PRIOR FILING DATE: 1999-12-09

; PRIOR APPLICATION NUMBER: 09/458,477

; PRIOR FILING DATE: 1999-12-09

; PRIOR APPLICATION NUMBER: 09/457,556

; PRIOR FILING DATE: 1999-12-09

; PRIOR APPLICATION NUMBER: 09/444,338

; PRIOR FILING DATE: 1999-11-19

; NUMBER OF SEQ ID NOS: 46

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 11

; LENGTH: 678

; TYPE: DNA

; ORGANISM: Discosoma species

US-10-006-922-11

Query Match 100.0%; Score 678; DB 13; Length 678;

Best Local Similarity 100.0%; Pred. No. 5.9e-216;

Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGGTCTTCCAAAGATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGATGGAAGGA 60  
|||||  
Db 1 ATGAGGTCTTCCAAAGATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGATGGAAGGA 60

QY 61 ACGGTCAATGGCACGAGTTTGAAATAGAAAGGCGAAGGAGAGGGGAGGCATACGAAGGC 120  
|||||  
Db 61 ACGGTCAATGGCACGAGTTTGAAATAGAAAGGCGAAGGAGAGGGGAGGCATACGAAGGC 120

QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTTGCCATTTTGGGATATT 180  
|||||  
Db 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTTGCCATTTTGGGATATT 180

QY 181 TTGTCAACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240  
|||||  
Db 181 TTGTCAACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240

QY 241 GACTATAAAAAGCTGTCATTTCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300  
|||||  
Db 241 GACTATAAAAAGCTGTCATTTCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300

QY 301 GACGGTGGCGTCGTACTGTAAACCCAGGATTCAGTTTGAGGATGGCTTTCATCTAC 360  
|||||  
Db 301 GACGGTGGCGTCGTACTGTAAACCCAGGATTCAGTTTGAGGATGGCTTTCATCTAC 360

QY 361 AAGGTCAAGTTTCAATGGCGTGAACCTTTCCCTCCGATGGACCTGTTATGCAAAAAGAGACA 420  
|||||  
Db 361 AAGGTCAAGTTTCAATGGCGTGAACCTTTCCCTCCGATGGACCTGTTATGCAAAAAGAGACA 420

QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
|||||  
Db 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGATCCTCGTGATGGCGTGTGAAAGGAGAG 480

QY 481 ATTCTAAGGCTCTGAAGCTGAAAGACGGTGGTCATTAACCTAGTTGAAATCAAAAGTATT 540  
|||||  
Db 481 ATTCTAAGGCTCTGAAGCTGAAAGACGGTGGTCATTAACCTAGTTGAAATCAAAAGTATT 540

QY 541 TACATGGCAAAAGACCTGTGCAGCTACAGGGTACTACTATGTTGACTCCAAACTGGAT 600  
|||||  
Db 541 TACATGGCAAAAGACCTGTGCAGCTACAGGGTACTACTATGTTGACTCCAAACTGGAT 600

QY 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC 660  
|||||  
Db 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC 660

QY 661 CACCATCTGTTCCCTTTAA 678  
|||||  
Db 661 CACCATCTGTTCCCTTTAA 678

RESULT 3

US-10-121-258-2

; Sequence 2, Application US/10121258

; Publication No. US20030059835A1

; GENERAL INFORMATION:

; APPLICANT: Tsien, Roger

; APPLICANT: Campbell, Robert

Handwritten signature and initials.

```

; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CP1
; CURRENT APPLICATION NUMBER: US/10/121,258
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Discosoma sp.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(678)
; OTHER INFORMATION: wild-type DsRed
US-10-121-258-2
```

Query Match 100.0%; Score 678; DB 14; Length 678;  
Best Local Similarity 100.0%; Pred. No. 5.9e-216;  
Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGGA 60
Db 1 ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGGA 60

QY 61 ACGGTCAATGGGCACGAGTTTGAATAGAACGCAAGGAGAGGGGAGGCCATACGAAGGC 120
Db 61 ACGGTCAATGGGCACGAGTTTGAATAGAACGCAAGGAGAGGGGAGGCCATACGAAGGC 120

QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGSGGGGACCTTTTGCCATTTCGTTGGGATATT 180
Db 121 CACAATACCGTAAAGCTTAAGGTAACCAAGSGGGGACCTTTTGCCATTTCGTTGGGATATT 180

QY 181 TTGTCACCAACAATTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240
Db 181 TTGTCACCAACAATTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240

QY 241 GACTATAAAAAGCTGTCTATTCTGAAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300
Db 241 GACTATAAAAAGCTGTCTATTCTGAAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300

QY 301 GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAAGGATGGCTGTTTCATCTAC 360
Db 301 GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAAGGATGGCTGTTTCATCTAC 360

QY 361 AAGGTCAAGTTTCATTGGCGTGAACTTTCCCTCCGATGGACCTGTTTATGCAAAAAGAGACA 420
Db 361 AAGGTCAAGTTTCATTGGCGTGAACTTTCCCTCCGATGGACCTGTTTATGCAAAAAGAGACA 420

QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGATCCTCGTGATGGCGTGTGAAAGGAGAG 480
Db 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGATCCTCGTGATGGCGTGTGAAAGGAGAG 480

QY 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATCAAAAAGTATT 540
Db 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATCAAAAAGTATT 540

QY 541 TACATGGCAAAAGAGCCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAT 600
Db 541 TACATGGCAAAAGAGCCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAT 600

QY 601 ATAACAAGCCACACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCGGAGCGC 660
Db 601 ATAACAAGCCACACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCGGAGCGC 660

QY 661 CACCATCTGTTTCCTTTAA 678
Db 661 CACCATCTGTTTCCTTTAA 678
```

```

RESULT 4
US-10-132-067-3
; Sequence 3, Application US/10132067
; Publication No. US20030203355A1
; GENERAL INFORMATION:
; APPLICANT: Bradbury, Andrew
; APPLICANT: Zeytun, Ahmet
; APPLICANT: Waldo, Geoffrey
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Fluorobodies: Binding Ligands With Intrinsic
; TITLE OF INVENTION: Fluorescence
; FILE REFERENCE: 021362-000600US
; CURRENT APPLICATION NUMBER: US/10/132,067
; CURRENT FILING DATE: 2002-04-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Discosoma sp.
; FEATURE:
; OTHER INFORMATION: red fluorescent protein (dsRED)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(678)
; OTHER INFORMATION: dsRED
US-10-132-067-3
```

Query Match 100.0%; Score 678; DB 17; Length 678;  
Best Local Similarity 100.0%; Pred. No. 5.9e-216;  
Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGGA 60
Db 1 ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGGA 60

QY 61 ACGGTCAATGGGCACGAGTTTGAATAGAACGCAAGGAGAGGGGAGGCCATACGAAGGC 120
Db 61 ACGGTCAATGGGCACGAGTTTGAATAGAACGCAAGGAGAGGGGAGGCCATACGAAGGC 120

QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGSGGGGACCTTTGCCATTTCGTTGGGATATT 180
Db 121 CACAATACCGTAAAGCTTAAGGTAACCAAGSGGGGACCTTTGCCATTTCGTTGGGATATT 180

QY 181 TTGTCACCAACAATTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240
Db 181 TTGTCACCAACAATTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240

QY 241 GACTATAAAAAGCTGTCTATTCTGAAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300
Db 241 GACTATAAAAAGCTGTCTATTCTGAAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300

QY 301 GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAAGGATGGCTGTTTCATCTAC 360
Db 301 GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAAGGATGGCTGTTTCATCTAC 360

QY 361 AAGGTCAAGTTTCATTGGCGTGAACTTTCCCTCCGATGGACCTGTTTATGCAAAAAGAGACA 420
Db 361 AAGGTCAAGTTTCATTGGCGTGAACTTTCCCTCCGATGGACCTGTTTATGCAAAAAGAGACA 420

QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGATCCTCGTGATGGCGTGTGAAAGGAGAG 480
Db 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGATCCTCGTGATGGCGTGTGAAAGGAGAG 480

QY 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATCAAAAAGTATT 540
Db 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATCAAAAAGTATT 540

QY 541 TACATGGCAAAAGAGCCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAT 600
Db 541 TACATGGCAAAAGAGCCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAT 600
```

*Hand date*

QY 601 ATAACAAGCCACAAGAACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC 660  
Db 601 ATAACAAGCCACAAGAACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC 660  
QY 661 CACCATCTGTTCTTTAA 678  
Db 661 CACCATCTGTTCTTTAA 678

RESULT 5  
US-10-335-517-6  
; Sequence 6, Application US/10335517  
; Publication No. US20030207248A1  
; GENERAL INFORMATION:  
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: TSUEN, Roger  
; APPLICANT: GONZALEZ, Jesus  
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY OPTICAL METHODS  
; FILE REFERENCE: REGEN1290-5  
; CURRENT APPLICATION NUMBER: US/10/335,517  
; CURRENT FILING DATE: 2002-12-31  
; PRIOR APPLICATION NUMBER: US/09/967,772  
; PRIOR FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: US 09/459,956  
; PRIOR FILING DATE: 1999-12-13  
; PRIOR APPLICATION NUMBER: US 08/765,860  
; PRIOR FILING DATE: 1996-12-19  
; PRIOR APPLICATION NUMBER: PCT/ US96/09652  
; PRIOR FILING DATE: 1996-06-06  
; PRIOR APPLICATION NUMBER: US 08/481,977  
; PRIOR FILING DATE: 1995-06-07  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 6  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Discosoma sp "red"  
US-10-335-517-6

Query Match 100.0%; Score 678; DB 17; Length 678;  
Best Local Similarity 100.0%; Pred. No. 5.9e-216;  
Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTTCAAGGTTTAAAGTTTCGCATGGAAGGA 60  
Db 1 ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTTCAAGGTTTAAAGTTTCGCATGGAAGGA 60  
QY 61 ACGGTCAATGGGCACGAGTTTGAATAGAACGAGGAGAGGGAGGCCATACGAAGGC 120  
Db 61 ACGGTCAATGGGCACGAGTTTGAATAGAACGAGGAGAGGGAGGCCATACGAAGGC 120  
QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
Db 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
QY 181 TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA 240  
Db 181 TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA 240  
QY 241 GACTATAAAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300  
Db 241 GACTATAAAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300  
QY 301 GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360  
Db 301 GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360  
QY 361 AAGGTCAAGTTTCATTTGGCGTGAACCTTTCCCTCCGATGGACCTGTTATGCAAAAGAAAGACA 420  
Db 361 AAGGTCAAGTTTCATTTGGCGTGAACCTTTCCCTCCGATGGACCTGTTATGCAAAAGAAAGACA 420  
QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTTTGAAGAGGAGAG 480

Db 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
QY 481 ATTCTAAGGCTCTGAAGCTGAAAGACGGTGGTTCATTACCTAGTTGAATTTCAAAGATATT 540  
Db 481 ATTCTAAGGCTCTGAAGCTGAAAGACGGTGGTTCATTACCTAGTTGAATTTCAAAGATATT 540  
QY 541 TACATGGCAAAAGAACGCTGTGACGTACCAAGGTACTACTATGTTGACTCCAAACTGGAT 600  
Db 541 TACATGGCAAAAGAACGCTGTGACGTACCAAGGTACTACTATGTTGACTCCAAACTGGAT 600  
QY 601 ATAACAAGCCACAAGAACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC 660  
Db 601 ATAACAAGCCACAAGAACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC 660  
QY 661 CACCATCTGTTCTTTAA 678  
Db 661 CACCATCTGTTCTTTAA 678

RESULT 6  
US-10-334-288-6  
; Sequence 6, Application US/10334288  
; Publication No. US20040002123A1  
; GENERAL INFORMATION:  
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: TSUEN, Roger  
; APPLICANT: GONZALEZ, Jesus  
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY OPTICAL METHODS  
; FILE REFERENCE: REGEN1290-5  
; CURRENT APPLICATION NUMBER: US/10/334,288  
; CURRENT FILING DATE: 2002-12-31  
; PRIOR APPLICATION NUMBER: US/09/967,772  
; PRIOR FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: US 09/459,956  
; PRIOR FILING DATE: 1999-12-13  
; PRIOR APPLICATION NUMBER: US 08/765,860  
; PRIOR FILING DATE: 1996-12-19  
; PRIOR APPLICATION NUMBER: PCT/ US96/09652  
; PRIOR FILING DATE: 1996-06-06  
; PRIOR APPLICATION NUMBER: US 08/481,977  
; PRIOR FILING DATE: 1995-06-07  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 6  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Discosoma sp "red"  
US-10-334-288-6

Query Match 100.0%; Score 678; DB 17; Length 678;  
Best Local Similarity 100.0%; Pred. No. 5.9e-216;  
Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTTCAAGGTTTAAAGTTTCGCATGGAAGGA 60  
Db 1 ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTTCAAGGTTTAAAGTTTCGCATGGAAGGA 60  
QY 61 ACGGTCAATGGGCACGAGTTTGAATAGAACGAGGAGAGGGAGGCCATACGAAGGC 120  
Db 61 ACGGTCAATGGGCACGAGTTTGAATAGAACGAGGAGAGGGAGGCCATACGAAGGC 120  
QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
Db 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
QY 181 TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA 240  
Db 181 TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA 240  
QY 241 GACTATAAAAAGCTGTCAATTTCCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300  
Db 241 GACTATAAAAAGCTGTCAATTTCCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300



QY 301 GACGGTGGCGTCGTTACTGTAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360  
Db 301 GACGGTGGCGTCGTTACTGTAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360  
QY 361 AAGGTCAAGTTTCATTGGCGTGAACCTTTCCTTCCGATGGACCTGTTATGCAAAAGAAGACA 420  
Db 361 AAGGTCAAGTTTCATTGGCGTGAACCTTTCCTTCCGATGGACCTGTTATGCAAAAGAAGACA 420  
QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
Db 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
QY 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT 540  
Db 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT 540  
QY 541 TACATGGCAAGAAGCCCTGTGTCAGCTACCAGGGTACTATGTTGACTCCAAACTGGAT 600  
Db 541 TACATGGCAAGAAGCCCTGTGTCAGCTACCAGGGTACTATGTTGACTCCAAACTGGAT 600  
QY 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGACGC 660  
Db 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGACGC 660  
QY 661 CACCATCTGTTCCCTTTAA 678  
Db 661 CACCATCTGTTCCCTTTAA 678

RESULT 7

US-10-311-030-5  
; Sequence 5, Application US/10311030  
; Publication No. US20040171107A1  
; GENERAL INFORMATION:  
; APPLICANT: Nelson, David  
; APPLICANT: Zamiaara, Elize  
; APPLICANT: Tsien, Roger  
; TITLE OF INVENTION: MODIFIED FLUORESCENT PROTEINS  
; FILE REFERENCE: 15916-032US1  
; CURRENT APPLICATION NUMBER: US/10/311,030  
; CURRENT FILING DATE: 2002-12-10  
; PRIOR APPLICATION NUMBER: PCT/US01/04625  
; PRIOR FILING DATE: 2001-02-13  
; PRIOR APPLICATION NUMBER: US 60/184,732  
; PRIOR FILING DATE: 2000-02-23  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Discosoma sp.

Query Match 100.0%; Score 678; DB 19; Length 678;  
Best Local Similarity 100.0%; Pred. No. 5.9e-216;  
Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGGTCTTCCAAGATGTTATCAAGGAGTTTCATCAGGTTTAAGGTTTCGCATGGAAGGA 60  
Db 1 ATGAGGTCTTCCAAGATGTTATCAAGGAGTTTCATCAGGTTTAAGGTTTCGCATGGAAGGA 60  
QY 61 ACGGTCAATGGGCACGAGTTTGAATAGAGCGGAAGAGAGAGGGGCCATACGAAGGC 120  
Db 61 ACGGTCAATGGGCACGAGTTTGAATAGAGCGGAAGAGAGAGGGGCCATACGAAGGC 120  
QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
Db 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
QY 181 TTGTCACCAACAATTTTCAGTATGGAAGCAAGGTATATGTCGAAGCACCCCTGCCGACATACCA 240  
Db 181 TTGTCACCAACAATTTTCAGTATGGAAGCAAGGTATATGTCGAAGCACCCCTGCCGACATACCA 240

QY 241 GACTATAAAAAGCTGTCTATTTCTGAAGGATTTAAATGGAAAGGTCATGAACCTTTGAA 300  
Db 241 GACTATAAAAAGCTGTCTATTTCTGAAGGATTTAAATGGAAAGGTCATGAACCTTTGAA 300  
QY 301 GACGGTGGCGTCGTTACTGTAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360  
Db 301 GACGGTGGCGTCGTTACTGTAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360  
QY 361 AAGGTCAAGTTTCATTGGCGTGAACCTTTCCTTCCGATGGACCTGTTATGCAAAAGAAGACA 420  
Db 361 AAGGTCAAGTTTCATTGGCGTGAACCTTTCCTTCCGATGGACCTGTTATGCAAAAGAAGACA 420  
QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
Db 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
QY 481 ATTCATAAGGCTCTGAAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT 540  
Db 481 ATTCATAAGGCTCTGAAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT 540  
QY 541 TACATGGCAAGAAGCCCTGTGTCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600  
Db 541 TACATGGCAAGAAGCCCTGTGTCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600  
QY 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGACGC 660  
Db 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGACGC 660  
QY 661 CACCATCTGTTCCCTTTAA 678  
Db 661 CACCATCTGTTCCCTTTAA 678

RESULT 8

US-10-656-029-21  
; Sequence 21, Application US/10656029  
; Publication No. US20050003367A1  
; GENERAL INFORMATION:  
; APPLICANT: VERTEX PHARMACEUTICALS INC.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR RAPID DEVELOPMENT OF  
; TITLE OF INVENTION: SCREENING ASSAYS  
; FILE REFERENCE: VPI/02-143WO2  
; CURRENT APPLICATION NUMBER: US/10/656,029  
; CURRENT FILING DATE: 2003-09-05  
; PRIOR APPLICATION NUMBER: 60/408,297  
; PRIOR FILING DATE: 2002-09-05  
; NUMBER OF SEQ ID NOS: 86  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 21  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Discosoma sp.  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(678)  
; OTHER INFORMATION: fluorescent protein  
US-10-656-029-21

Query Match 100.0%; Score 678; DB 21; Length 678;  
Best Local Similarity 100.0%; Pred. No. 5.9e-216;  
Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGGTCTTCCAAGATGTTATCAAGGAGTTTCATCAGGTTTAAGGTTTCGCATGGAAGGA 60  
Db 1 ATGAGGTCTTCCAAGATGTTATCAAGGAGTTTCATCAGGTTTAAGGTTTCGCATGGAAGGA 60  
QY 61 ACGGTCAATGGGCACGAGTTTGAATAGAGCGGAAGGAGAGGGGCCATACGAAGGC 120  
Db 61 ACGGTCAATGGGCACGAGTTTGAATAGAGCGGAAGGAGAGGGGCCATACGAAGGC 120  
QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
Db 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180





QY	61	ACGGTCAATGGGCACGAGTTTGAATAAGAAAGCGGAGAGAGGGAGGCCATACGAAGGC	120
DB	114	ACGGTCAATGGGCACGAGTTTGAATAAGAAAGCGGAGAGGGAGGCCATACGAAGGC	173
QY	121	CACAATACCGTAAAGCTTAAGSTAACCAAGGGGGGACCTTTGCCATTTGCTTGGGATATT	180
DB	174	CACAATACCGTAAAGCTTAAGSTAACCAAGGGGGGACCTTTGCCATTTGCTTGGGATATT	233
QY	181	TTGTCAACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA	240
DB	234	TTGTCAACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA	293
QY	241	GACTATAAAAAGCTGTCATTTCTCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTTGAA	300
DB	294	GACTATAAAAAGCTGTCATTTCTCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTTGAA	353
QY	301	GACGGTGGCGTCGTTACTGTAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC	360
DB	354	GACGGTGGCGTCGTTACTGTAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC	413
QY	361	AAGGTCGAAGTTTCATTGGCGTGAACCTTTCTCCGATGGACCTGTTATGCAAAAAGAACA	420
DB	414	AAGGTCGAAGTTTCATTGGCGTGAACCTTTCTCCGATGGACCTGTTATGCAAAAAGAACA	473
QY	421	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG	480
DB	474	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG	533
QY	481	ATTATAAGGCTCTGAAGCTGAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT	540
DB	534	ATTATAAGGCTCTGAAGCTGAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT	593
QY	541	TACATGGCAAGAAGCCCTGTGCAGCTACCAGGTACTACTATGTTGACTCCAAACTGGAT	600
DB	594	TACATGGCAAGAAGCCCTGTGCAGCTACCAGGTACTACTATGTTGACTCCAAACTGGAT	653
QY	601	ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCGAGGGACGC	660
DB	654	ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCGAGGGACGC	713
QY	661	CACCATCTGTTCCCTTTAA	678
DB	714	CACCATCTGTTCCCTTTAA	731

## RESULT 11

```

US-09-794-308-11
; Sequence 11, Application US/09794308
; Publication No. US20030170911A1
; GENERAL INFORMATION:
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: TS'EN, Roger
; APPLICANT: ZACHARIAS, David
; APPLICANT: BAIRD, Geoffrey
; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT
; FILE REFERENCE: REGEN1530
; CURRENT APPLICATION NUMBER: US/09/794,308
; CURRENT FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 859
; TYPE: DNA
; ORGANISM: Discosoma sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)..(731)
US-09-794-308-11

```

QY	1	ATGAGGCTCTTCCAAGAAATGTTATCAAGGAGTTTCATAGGTTTAAAGTTTCGCATGGAAGGA	60
DB	54	ATGAGGCTCTTCCAAGAAATGTTATCAAGGAGTTTCATAGGTTTAAAGTTTCGCATGGAAGGA	113
QY	61	ACGGTCAATGGGCACGAGTTTGAAATAGAAGGCGAAGGAGGGGAGGCCATACGAAGGC	120
DB	114	ACGGTCAATGGGCACGAGTTTGAAATAGAAGGCGAAGGAGGGGAGGCCATACGAAGGC	173
QY	121	CACAAATACCGTAAAGCTTTAAGGTAAACCAAGGGGGGACCTTTGCCAATTTGCTTTGGGATATT	180
DB	174	CACAAATACCGTAAAGCTTTAAGGTAAACCAAGGGGGGACCTTTGCCAATTTGCTTTGGGATATT	233
QY	181	TTGTCACCACAAATTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA	240
DB	234	TTGTCACCACAAATTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA	293
QY	241	GACTATAAAAAGCTGTCTATTCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA	300
DB	294	GACTATAAAAAGCTGTCTATTCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA	353
QY	301	GACGGTGGCGTCGTTACTGTAACCCAGGATTCAGTTTGACGATGGCTGTTTCATCTAC	360
DB	354	GACGGTGGCGTCGTTACTGTAACCCAGGATTCAGTTTGACGATGGCTGTTTCATCTAC	413
QY	361	AAGGTCAAGTTTCATTTGGCGTGAACTTTCTTCCGATGGACCTGTATGCAAAAAGAACA	420
DB	414	AAGGTCAAGTTTCATTTGGCGTGAACTTTCTTCCGATGGACCTGTATGCAAAAAGAACA	473
QY	421	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGATCCTCGTATGGCGTGTTGAAAGGAGAG	480
DB	474	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGATCCTCGTATGGCGTGTTGAAAGGAGAG	533
QY	481	ATTCTAAGGCTCTGAAGCTGAAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAATTT	540
DB	534	ATTCTAAGGCTCTGAAGCTGAAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAATTT	593
QY	541	TACATGGCAAAGAAGCCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT	600
DB	594	TACATGGCAAAGAAGCCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT	653
QY	601	ATAACAAGCCACAAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC	660
DB	654	ATAACAAGCCACAAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC	713
QY	661	CACCATCTGTTCCCTTTAA	678
DB	714	CACCATCTGTTCCCTTTAA	731

RESULT 12

```

US-09-865-291-11
; Sequence 11, Application US/09865291
; Publication No. US20030186229A1
; GENERAL INFORMATION:
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: TSIEN, Roger
; APPLICANT: TING, Alice
; APPLICANT: ZHANG, Jin
; TITLE OF INVENTION: EMISSION RATIONETRIC INDICATORS OF PHOSPHORYLATION
; FILE REFERENCE: REGEN1550
; CURRENT APPLICATION NUMBER: US/09/865,291
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 859
; TYPE: DNA
; ORGANISM: Discosoma sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)..(731)
US-09-865-291-11

```

```

Query Match      100.0%; Score 678; DB 10; Length 859;
Best Local Similarity 100.0%; Pred. No. 6.7e-216;
Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
Query Match      100.0%;   Score 678;   DB 10;   Length 859;
Best Local Similarity 100.0%;   Pred. No. 6.7e-216;
Matches 678;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

QY      1  ATGAGGCTCTCCAGAAGTGTATCAAGGAGTTTCATGAGGTTTAAGGTTCCGATGGAAGGA 60
      |||||||
Db      54  ATGAGGCTCTCCAGAAGTGTATCAAGGAGTTTCATGAGGTTTAAGGTTCCGATGGAAGGA 113

QY      61  ACGGTCAATGGGCACGAGTTTGAATAGAAAGGCGAAGGAGAGGGGAGGCATACGAAGGC 120
      |||||||
Db     114  ACGGTCAATGGGCACGAGTTTGAATAGAAAGGCGAAGGAGAGGGGAGGCATACGAAGGC 173

QY     121  CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTGTCTGGGATATT 180
      |||||||
Db     174  CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTGTCTGGGATATT 233

QY     181  TTGTCACCACAATTTCAAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240
      |||||||
Db     234  TTGTCACCACAATTTCAAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 293

QY     241  GACTATAAAAAGCTGTCATTTCCCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAA 300
      |||||||
Db     294  GACTATAAAAAGCTGTCATTTCCCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAA 353

QY     301  GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360
      |||||||
Db     354  GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 413

QY     361  AAGGTCAAGTTCATTTGGCGTGAACTTTCCTTCGGATGGACCTGTTATGCAAAAAGAGACA 420
      |||||||
Db     414  AAGGTCAAGTTCATTTGGCGTGAACTTTCCTTCGGATGGACCTGTTATGCAAAAAGAGACA 473

QY     421  ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTATGGCGTGTGAAAAGGAGAG 480
      |||||||
Db     474  ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTATGGCGTGTGAAAAGGAGAG 533

QY     481  ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT 540
      |||||||
Db     534  ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT 593

QY     541  TACATGGCAAAGAACCTGTGCAGCTACCAGGCTACTACTATGTTGACTCCAAACTGGAT 600
      |||||||
Db     594  TACATGGCAAAGAACCTGTGCAGCTACCAGGCTACTACTATGTTGACTCCAAACTGGAT 653

QY     601  ATAACAAGCCACAAAGAACTATACAATCGTTGAGCAGTATGAAAGAACCGGAGGCACGC 660
      |||||||
Db     654  ATAACAAGCCACAAAGAACTATACAATCGTTGAGCAGTATGAAAGAACCGGAGGCACGC 713

QY     661  CACCATCTGTTCCCTTTAA 678
      |||||||
Db     714  CACCATCTGTTCCCTTTAA 731
```

```
RESULT 13
US-10-433-640-12
; Sequence 12, Application US/10433640
; Publication No. US20040115792A1
; GENERAL INFORMATION:
; APPLICANT: Lichtenberg-Frate, Hella
; TITLE OF INVENTION: YEAST STRAIN FOR TESTING THE GENOTOXICITY AND CYTOTOXICITY OF
; FILE OF INVENTION: COMPLEX ENVIRONMENTAL CONTAMINATIONS
; FILE REFERENCE: 1487/3
; CURRENT APPLICATION NUMBER: US/10/433,640
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/14610
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: DE 10061872.3
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12
; LENGTH: 859
,
```

```
; TYPE: DNA
; ORGANISM: Discosoma sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)..(731)
US-10-433-640-12

Query Match      100.0%;   Score 678;   DB 19;   Length 859;
Best Local Similarity 100.0%;   Pred. No. 6.7e-216;
Matches 678;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

QY      1  ATGAGTCTTCCAAAGAATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGGA 60
      |||||||
Db      54  ATGAGTCTTCCAAAGAATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGGA 113

QY      61  ACGGTCAATGGGCACGAGTTTGAATAGAAAGGCGAAGGAGAGGGGAGGCCATACGAAGGC 120
      |||||||
Db     114  ACGGTCAATGGGCACGAGTTTGAATAGAAAGGCGAAGGAGAGGGGAGGCCATACGAAGGC 173

QY     121  CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTGCTTGGGATATT 180
      |||||||
Db     174  CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTGCTTGGGATATT 233

QY     181  TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240
      |||||||
Db     234  TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 293

QY     241  GACTATAAAAAGCTGTCAATTTCCCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTTGA 300
      |||||||
Db     294  GACTATAAAAAGCTGTCAATTTCCCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTTGA 353

QY     301  GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360
      |||||||
Db     354  GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 413

QY     361  AAGGTCAAGTTCATTTGGCGTGAACTTTCCTTCGGATGGACCTTATGCAAAAAGAGACA 420
      |||||||
Db     414  AAGGTCAAGTTCATTTGGCGTGAACTTTCCTTCGGATGGACCTTATGCAAAAAGAGACA 473

QY     421  ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTATGGCGTGTGAAAAGGAGAG 480
      |||||||
Db     474  ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTATGGCGTGTGAAAAGGAGAG 533

QY     481  ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT 540
      |||||||
Db     534  ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT 593

QY     541  TACATGGCAAAGAACCTGTGCAGCTACCAGGCTACTACTATGTTGACTCCAAACTGGAT 600
      |||||||
Db     594  TACATGGCAAAGAACCTGTGCAGCTACCAGGCTACTACTATGTTGACTCCAAACTGGAT 653

QY     601  ATAACAAGCCACAAAGAACTATACAATCGTTGAGCAGTATGAAAGAACCGGAGGCACGC 660
      |||||||
Db     654  ATAACAAGCCACAAAGAACTATACAATCGTTGAGCAGTATGAAAGAACCGGAGGCACGC 713

QY     661  CACCATCTGTTCCCTTTAA 678
      |||||||
Db     714  CACCATCTGTTCCCTTTAA 731

RESULT 14
US-10-885-988-11
; Sequence 11, Application US/10885988
; Publication No. US20040259165A1
; GENERAL INFORMATION:
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: TSJEN, Roger
; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS
; FILE REFERENCE: REGEN1530-2
; CURRENT APPLICATION NUMBER: US/10/885,988
; CURRENT FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: US/09/866,538
```



; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 859
; TYPE: DNA
; ORGANISM: Discosoma sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)..(731)
US-10-885-988-11

Query Match 100.0%; Score 678; DB 20; Length 859;
Best Local Similarity 100.0%; Pred. No. 6.7e-216;
Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGGTCTTCCAAGATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGGA 60
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
54 ATGAGGTCTTCCAAGATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGGA 113
QY 61 ACGGTCAATGGGCACGAGTTTGAATAGAGCGGAAGGAGAGGGGAGGCCATACGAAGGC 120
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
114 ACGGTCAATGGGCACGAGTTTGAATAGAGCGGAAGGAGAGGGGAGGCCATACGAAGGC 173
QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
174 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 233
QY 181 TTGTCACCACAATTTTCAGTATGGAAGCAAGGTAATATGTCGAAGCACCTTGCCGACATACCA 240
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
234 TTGTCACCACAATTTTCAGTATGGAAGCAAGGTAATATGTCGAAGCACCTTGCCGACATACCA 293
QY 241 GACTATAAAAAGCTGTCAATTTCTGAAGGATTTAAATGGGAAAGGGTCAATGAATTTGAA 300
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
294 GACTATAAAAAGCTGTCAATTTCTGAAGGATTTAAATGGGAAAGGGTCAATGAATTTGAA 353
QY 301 GACGGTGGCGTCTGTTACTGTAAACCCAGGATTCAGTTTGCAAGGATGGCTGTTTCATCTAC 360
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
354 GACGGTGGCGTCTGTTACTGTAAACCCAGGATTCAGTTTGCAAGGATGGCTGTTTCATCTAC 413
QY 361 AAGGTCAAGTTTCATTTGGCGTGAACTTTCCCTCCGATGGACCTGTTATGCAAAAGAGACA 420
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
414 AAGGTCAAGTTTCATTTGGCGTGAACTTTCCCTCCGATGGACCTGTTATGCAAAAGAGACA 473
QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGTATGGCGTGTGAAAGGAGAG 480
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
474 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGTATGGCGTGTGAAAGGAGAG 533
QY 481 ATTATAAGGCTCTGAAGCTGAAAGACGGTGGTCAATTAACCTAGTTGAATTCAAAAGTATT 540
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
534 ATTATAAGGCTCTGAAGCTGAAAGACGGTGGTCAATTAACCTAGTTGAATTCAAAAGTATT 593
QY 541 TACATGGCAAAAGAGCCCTGTGCAGCTACCAAGGTTACTATGTTGACTCCAAACTGGAT 600
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
594 TACATGGCAAAAGAGCCCTGTGCAGCTACCAAGGTTACTATGTTGACTCCAAACTGGAT 653
QY 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCCGAGGGACGC 660
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
654 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCCGAGGGACGC 713
QY 661 CACCATCTGTTCCCTTTAA 678
Db |||||||||||||||||||
714 CACCATCTGTTCCCTTTAA 731

RESULT 15
US-10-857-622-11
; Sequence 11, Application US/10857622
; Publication No. US20050026234A1
; GENERAL INFORMATION:
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: VIOLIN, Jonathan
; APPLICANT: NEWTON, Alexandra

; APPLICANT: TSJEN, Roger
; APPLICANT: ZHANG, Jin
; TITLE OF INVENTION: Emission Ratiometric Indicators of Phosphorylation By C-Kinase
; FILE REFERENCE: 39754-0891 CPC1CP2
; CURRENT APPLICATION NUMBER: US/10/857,622
; CURRENT FILING DATE: 2004-05-28
; PRIOR APPLICATION NUMBER: US 09/865,291
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/396,003
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: US 08/792,553
; PRIOR FILING DATE: 1997-01-31
; PRIOR APPLICATION NUMBER: US 594,575
; PRIOR FILING DATE: 1996-01-31
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 859
; TYPE: DNA
; ORGANISM: Discosoma sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)..(731)
US-10-857-622-11

Query Match 100.0%; Score 678; DB 21; Length 859;
Best Local Similarity 100.0%; Pred. No. 6.7e-216;
Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGGTCTTCCAAGATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGGA 60
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
54 ATGAGGTCTTCCAAGATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGGA 113
QY 61 ACGGTCAATGGGCACGAGTTTGAATAGAGCGGAAGGAGAGGGGAGGCCATACGAAGGC 120
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
114 ACGGTCAATGGGCACGAGTTTGAATAGAGCGGAAGGAGAGGGGAGGCCATACGAAGGC 173
QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGGACCTTTGCCATTTGCTTGGGATATT 180
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
174 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGGACCTTTGCCATTTGCTTGGGATATT 233
QY 181 TTGTCACCACAATTTTCAGTATGGAAGCAAGGTAATATGTCGAAGCACCTTGCCGACATACCA 240
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
234 TTGTCACCACAATTTTCAGTATGGAAGCAAGGTAATATGTCGAAGCACCTTGCCGACATACCA 293
QY 241 GACTATAAAAAGCTGTCAATTTCTGAAGGATTTAAATGGGAAAGGGTCAATGAATTTGAA 300
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
294 GACTATAAAAAGCTGTCAATTTCTGAAGGATTTAAATGGGAAAGGGTCAATGAATTTGAA 353
QY 301 GACGGTGGCGTCTGTTACTGTAAACCCAGGATTCAGTTTGCAAGGATGGCTGTTTCATCTAC 360
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
354 GACGGTGGCGTCTGTTACTGTAAACCCAGGATTCAGTTTGCAAGGATGGCTGTTTCATCTAC 413
QY 361 AAGGTCAAGTTTCATTTGGCGTGAACTTTCCCTCCGATGGACCTGTTATGCAAAAGAGACA 420
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
414 AAGGTCAAGTTTCATTTGGCGTGAACTTTCCCTCCGATGGACCTGTTATGCAAAAGAGACA 473
QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGTATGGCGTGTGAAAGGAGAG 480
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
474 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGTATGGCGTGTGAAAGGAGAG 533
QY 481 ATTATAAGGCTCTGAAGCTGAAAGACGGTGGTCAATTAACCTAGTTGAATTCAAAAGTATT 540
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
534 ATTATAAGGCTCTGAAGCTGAAAGACGGTGGTCAATTAACCTAGTTGAATTCAAAAGTATT 593
QY 541 TACATGGCAAAAGAGCCCTGTGCAGCTACCAAGGTTACTATGTTGACTCCAAACTGGAT 600
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
594 TACATGGCAAAAGAGCCCTGTGCAGCTACCAAGGTTACTATGTTGACTCCAAACTGGAT 653
QY 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCCGAGGGACGC 660
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
654 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCCGAGGGACGC 713



Qy 661 CACCATCTGTTCCCTTAA 678  
| | | | | | | | | | | | | | | | | |  
Db 714 CACCATCTGTTCCCTTAA 731

Search completed: July 1, 2005, 18:32:40  
Job time : 573 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2005, 15:58:43 ; Search time 170 Seconds  
(without alignments)  
6525.857 Million cell updates/sec

Title: US-10-006-922A-11  
Perfect score: 678  
Sequence: 1 atgaggctctccaagaatgt.....gccaccatctgttccttaa 678

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	678	100.0	678	3	US-09-459-956-6
2	450.6	66.5	723	4	US-10-152-296-1
3	409.6	60.4	699	3	US-09-459-956-5
4	271	40.0	801	3	US-09-459-956-7
5	178.4	26.3	690	3	US-09-459-956-2
6	147.2	21.7	696	3	US-09-459-956-4
7	141.4	20.9	1482	4	US-09-977-897-1
8	133.4	19.7	696	3	US-09-459-956-3
9	133.2	19.6	1079	3	US-09-609-161B-15
10	133.2	19.6	1079	4	US-09-626-581D-64
11	133.2	19.6	1079	4	US-09-415-765B-64
12	133.2	19.6	1079	4	US-09-626-580C-64
13	133.2	19.6	1085	3	US-09-277-716-15
14	132.4	19.5	1104	3	US-09-277-716-30
15	132.4	19.5	1104	3	US-09-609-161B-30
16	132.4	19.5	1279	3	US-09-277-716-31
17	132.4	19.5	1279	3	US-09-609-161B-31
18	130	19.2	1021	4	US-09-839-650-2
19	128.4	18.9	720	4	US-09-839-650-1
20	42	6.2	7218	1	US-08-232-463-14
21	35.8	5.3	1141	4	US-09-806-708B-22
22	34.4	5.1	2748	4	US-09-540-236-1106
23	34.4	5.1	19988	4	US-09-596-002-10
24	34.2	5.0	14207	4	US-09-949-016-12187
25	34.2	5.0	14207	4	US-09-949-016-17460
26	33.6	5.0	591	4	US-09-248-796A-2585
27	33.2	4.9	601	4	US-09-949-016-43022

28	33.2	4.9	580073	4	US-08-545-528D-1	Sequence 1, Appli
29	32.4	4.8	601	4	US-09-949-016-45637	Sequence 45637, A
c 30	32.4	4.8	30032	4	US-09-949-016-13933	Sequence 13933, A
c 31	32.4	4.8	150394	4	US-09-949-016-13042	Sequence 13042, A
32	32.2	4.7	156942	4	US-09-949-016-12227	Sequence 12227, A
33	32.2	4.7	156950	4	US-09-949-016-15946	Sequence 15946, A
34	32	4.7	601	4	US-09-949-016-43023	Sequence 43023, A
35	32	4.7	1098	4	US-09-716-793A-3	Sequence 3, Appli
c 36	32	4.7	187916	4	US-09-949-016-12980	Sequence 12980, A
37	31.8	4.7	41736	4	US-09-949-016-17091	Sequence 17091, A
c 38	31.4	4.6	1452	4	US-09-614-221A-317	Sequence 317, App
c 39	31.4	4.6	14226	4	US-09-949-016-14002	Sequence 14002, A
40	31.2	4.6	2563	4	US-09-949-016-1360	Sequence 1360, Ap
41	31.2	4.6	2570	4	US-09-919-497-16	Sequence 16, Appl
42	31.2	4.6	4100	4	US-09-620-312D-81	Sequence 81, Appl
43	31.2	4.6	101349	4	US-09-949-016-17433	Sequence 17433, A
c 44	31.2	4.6	154023	4	US-09-949-016-17057	Sequence 17057, A
c 45	31.2	4.6	167708	4	US-09-949-016-16423	Sequence 16423, A

ALIGNMENTS

RESULT 1

US-09-459-956-6

; Sequence 6, Application US/09459956

; Patent No. 6342379

; GENERAL INFORMATION:

; APPLICANT: Tsien, Roger Y.

; APPLICANT: Gonzalez, III, Jesus E.

; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY

; TITLE OF INVENTION: OPTICAL METHODS

; FILE REFERENCE: REGEN1290-4

; CURRENT APPLICATION NUMBER: US/09/459,956

; CURRENT FILING DATE: 1999-12-13

; PRIOR APPLICATION NUMBER: 08/765,860

; PRIOR FILING DATE: 1999-05-08

; PRIOR APPLICATION NUMBER: 08/481,977

; PRIOR FILING DATE: 1995-06-07

; PRIOR APPLICATION NUMBER: PCT/US96/09652

; PRIOR FILING DATE: 1996-06-06

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 6

; LENGTH: 678

; TYPE: DNA

; ORGANISM: Discosoma sp

US-09-459-956-6

Query Match 100.0%; Score 678; DB 3; Length 678;  
Best Local Similarity 100.0%; Pred. No. 7.7e-229;  
Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTTCATGAGGTTTAAAGTTCCGATGGAAGGA	60
Db	1	ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTTCATGAGGTTTAAAGTTCCGATGGAAGGA	60
Qy	61	ACGGTCAATGGGCACGAGTTTGAATAGAACGCGAAGGAGAGGGAGGCCATACGAAGGC	120
Db	61	ACGGTCAATGGGCACGAGTTTGAATAGAACGCGAAGGAGAGGGAGGCCATACGAAGGC	120
Qy	121	CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT	180
Db	121	CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT	180
Qy	181	TTGTCAACCAATTTTCAGTATGGAAGCAAGGATATATGTCAAGCACCCCTCCGACATACCA	240
Db	181	TTGTCAACCAATTTTCAGTATGGAAGCAAGGATATATGTCAAGCACCCCTCCGACATACCA	240
Qy	241	GACTATAAAAGGTGTCTTCTGAAGGATTTAAATGGGAAAGGGTCAATGAACCTTTTGAA	300
Db	241	GACTATAAAAGGTGTCTTCTGAAGGATTTAAATGGGAAAGGGTCAATGAACCTTTTGAA	300

QY 301 GACGGTGGCGTCTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360  
Db |||||  
QY 301 GACGGTGGCGTCTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360  
Db |||||  
QY 361 AAGGTCAAGTTCAATTGGCGTGAACCTTTCTTCCGATGGACCTGTTATGCAAAAGAAGACA 420  
Db |||||  
QY 361 AAGGTCAAGTTCAATTGGCGTGAACCTTTCTTCCGATGGACCTGTTATGCAAAAGAAGACA 420  
Db |||||  
QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
Db |||||  
QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
Db |||||  
QY 481 ATTATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT 540  
Db |||||  
QY 481 ATTATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT 540  
Db |||||  
QY 541 TACATGGCAAAGAGCCTGTGCAGCTACCAGGTACTACTATGTTGACTCCAAAAGTGGAT 600  
Db |||||  
QY 541 TACATGGCAAAGAGCCTGTGCAGCTACCAGGTACTACTATGTTGACTCCAAAAGTGGAT 600  
Db |||||  
QY 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC 660  
Db |||||  
QY 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC 660  
Db |||||  
QY 661 CACCATCTGTTCCCTTTAA 678  
Db |||||  
QY 661 CACCATCTGTTCCCTTTAA 678  
Db |||||

RESULT 2  
US-10-152-296-1  
; Sequence 1, Application US/10152296  
; Patent No. 6723537  
; GENERAL INFORMATION:  
; APPLICANT: Peelle, Beau  
; APPLICANT: Rigel Pharmaceuticals, Incorporated  
; TITLE OF INVENTION: Directed Evolution of Protein in Mammalian Cells  
; FILE REFERENCE: 021044-000110US  
; CURRENT APPLICATION NUMBER: US/10/152,296  
; PRIOR FILING DATE: 2002-12-10  
; PRIOR APPLICATION NUMBER: US 60/291,871  
; PRIOR FILING DATE: 2001-05-18  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 723  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:mammalian  
; OTHER INFORMATION: codon-optimized variant (DsRED) of Discosoma sp.  
; OTHER INFORMATION: "red" red fluorescent protein (RFP)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(723)  
; OTHER INFORMATION: DsRED  
US-10-152-296-1

Query Match 66.5%; Score 450.6; DB 4; Length 723;  
Best Local Similarity 79.3%; Pred. No. 2e-148;  
Matches 534; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY 2 TGAGGTCTTCCAAGATGTTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGAA 61  
Db |||||  
QY 5 TGGGCTCTCTCCAAGAACGTCATCAAGGAGTTTCATGCGCTTCAAGGTGCGCATGGAGGGCA 64  
Db |||||  
QY 62 CGGTCAATGGGCACGAGTTTGAATAGAGCGGAAGAGAGGGGAGGGCCATACGAAGGCC 121  
Db |||||  
QY 65 CCGTGAACGGCCACGAGTTCGAGATCGAGGCGGAGGGCGAGGGCCGCCCTACGAGGGCC 124  
Db |||||  
QY 122 ACAATACCGTAAAGCTTAAGGTTAACCAGGGGGGACCTTTGCCATTTGCTTGGGATATTT 181  
Db |||||  
QY 125 ACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGGCCCCCTTGCCTTGGGACATCC 184  
Db |||||

QY 182 TGTCAACCAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCAG 241  
Db |||||  
QY 185 TGTCCCCCAGTTCAGTACGGTCCAAGGTGTACGTGAAGCAACCCCGCGACATCCCCG 244  
Db |||||  
QY 242 ACTATAAAAGCTGTTCATTTCTTCTGAAAGGATTTAAATGGGAAAGGTCATGAACTTTGAAG 301  
Db |||||  
QY 245 ACTACAAGAACTGTCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGG 304  
Db |||||  
QY 302 ACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTACA 361  
Db |||||  
QY 305 ACGGCGCGTGGTGACCGTGACCCAGGACTCCTCCCTGCAGGACGGCTGCTTCATCTACA 364  
Db |||||  
QY 362 AGGTCAAGTTCATTTGGCGTGAACCTTTCTTCCGATGGACCTGTTATGCAAAAGAAGACAA 421  
Db |||||  
QY 365 AGGTGAAGTTCATCGGCGTGAACCTTCCCTCCGACGGCCCCGTAATGCAGAAGAAGACCA 424  
Db |||||  
QY 422 TGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAGA 481  
Db |||||  
QY 425 TGGGCTGGGAGGCCCTCCACCGAGCGCTGTACCCCCCGCAGCGCGTGTGAAGGGCGAGA 484  
Db |||||  
QY 482 TTCATAAGGCTCTGAAGCTGAAAGACGGTGGTGCTATTACCTAGTTGAATTCAAAAGTATTT 541  
Db |||||  
QY 485 TCCACAAGGCCCTGAAGCTGAAAGGACGGCGGGCCACTACCTGGTGGAGTTCAAGAGTATCT 544  
Db |||||  
QY 542 ACATGGCAAAGAACGCTGTGCAGTACCAGGGTACTACTATGTTGACTCCAAAAGTGGATA 601  
Db |||||  
QY 545 ACATGGCCAAGAAAGCCCCGTGACGTGCCCCGGCTACTACTACGTGGACTCCAAGCTGGACA 604  
Db |||||  
QY 602 TAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCCGAGGGACGCC 661  
Db |||||  
QY 605 TCACCTCCCAACAACGAGGACTACACCATCGTGGAGCAGTACGAGCGCACCCGAGGGCCGCC 664  
Db |||||  
QY 662 ACCATCTGTTCTT 674  
Db |||||  
QY 665 ACCACCTGTTCTT 677  
Db |||||

RESULT 3  
US-09-459-956-5  
; Sequence 5, Application US/09459956  
; Patent No. 6342379  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger Y.  
; APPLICANT: Gonzalez, III, Jesus E.  
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY  
; TITLE OF INVENTION: OPTICAL METHODS  
; FILE REFERENCE: RGEN1290-4  
; CURRENT APPLICATION NUMBER: US/09/459,956  
; CURRENT FILING DATE: 1999-12-13  
; PRIOR APPLICATION NUMBER: 08/765,860  
; PRIOR FILING DATE: 1999-05-08  
; PRIOR APPLICATION NUMBER: 08/481,977  
; PRIOR FILING DATE: 1995-06-07  
; PRIOR APPLICATION NUMBER: PCT/US96/09652  
; PRIOR FILING DATE: 1996-06-06  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 699  
; TYPE: DNA  
; ORGANISM: Discosoma striata  
US-09-459-956-5

Query Match 60.4%; Score 409.6; DB 3; Length 699;  
Best Local Similarity 76.2%; Pred. No. 6e-134;  
Matches 518; Conservative 0; Mismatches 159; Indels 3; Gaps 1;

QY 1 ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTTCATGAGGTTTAAAGTTTCGCATGGAAGGA 60  
Db |||||  
QY 1 ATGAGTTGTTCCAAGAGTGTGATCAAGGAAGAAATGTTGATCGATCTTTCATCTGGAAGGA 60  
Db |||||  
QY 61 ACGGTCAATGGGCACGAGTTTGAATAGAAAGCGGAAGGAGAGGGGAGGCCATACGAAGGC 120  
Db |||||

Db 61 AGTTCAATGGGCACCTACTTTGAAATAAAAGGCAAGAAAGGACAGCCTAATGAAGGC 120  
Qy 121 CACAATACCGTAAAGCTTAAGGTAAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
Db 121 ACCAATACCGTCACGCTCGAGGTTACCAAGGGTGGACCTCTGCCATTTGGTGGCATAATT 180  
Qy 181 TTGTCAACCACAATTTCAAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240  
Db 181 TTGTGCCCACAATTTCAAGTATGGAAGCAAGGTATATGTCCACCACCCCTGCCAACAATACAT 240  
Qy 241 GACTATAAAAAGCTGTCAATTTCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300  
Db 241 GATTATCTAAAGCTGTCAATTTCCGGAGGGATATACATGGGAACGGTCCATGCACCTTTGAA 300  
Qy 301 GACGGTGGCGTCTACTGTAAACCCAGGATTCAGATTTCAGGATGGCTGTTTCTACTAC 360  
Db 301 GACGGTGGCTTGTGTGTATCACCAATGATATCAGTTTGACAGGCAACTGTTTCTACTAC 360  
Qy 361 AAGGTCAAGTTTCAATGGCGTGAACTTTCTCCGATGGACCTGTTTATGCAAAAGAAAGACA 420  
Db 361 GACATCAAGTTTCACTGGCTTGAACCTTTCTCCAAATGGACCCGTTGTGCAAGAAAGACA 420  
Qy 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
Db 421 ACTGGCTGGGAACCCAGCACTGAGCGTTTGTATCCTCGTGATGGTGTGTGATAGGAGAC 480  
Qy 481 ATTATAAGGCTCTGAAGCTGAAAGACGGTGGTCAATACCTAGTTGAAATTCAAAAGTATT 540  
Db 481 ATCCATCATGCTCTGACAGTTGAAGGAGGTGGTCAATACGCTGTGACATTAATAAAGTGT 540  
Qy 541 TACATGGCAAGAAGCCTG--TGCAGCTACCAGGCTACTACTATGTTGACTCCAAACTG 597  
Db 541 TACAGGGCCAAGAAGCGCCTTGAAGATGCCAGGGTATCACTATGTTGACACCAAACTG 600  
Qy 598 GATATAACAAGCCCAACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCGGAGGA 657  
Db 601 GTTATATGGAACAACGACAAAGAATTTCATGAAGTTGAGGAGCATGAATCGCCGTTGCA 660  
Qy 658 CGCCACCACCTGTTCTCTTTA 677  
Db 661 CGCCACCACCTCGTTCTATGA 680

RESULT 4  
US-09-459-956-7  
; Sequence 7, Application US/09459956  
; Patent No. 6342379  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger Y.  
; APPLICANT: Gonzalez, Iii, Jesus E.  
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY  
; TITLE OF INVENTION: OPTICAL METHODS  
; FILE REFERENCE: REGEN1290-4  
; CURRENT APPLICATION NUMBER: US/09/459,956  
; CURRENT FILING DATE: 1999-12-13  
; PRIOR APPLICATION NUMBER: 08/765,860  
; PRIOR FILING DATE: 1999-05-08  
; PRIOR APPLICATION NUMBER: 08/481,977  
; PRIOR FILING DATE: 1995-06-07  
; PRIOR APPLICATION NUMBER: PCT/US96/09652  
; PRIOR FILING DATE: 1996-06-06  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 7  
; TYPE: DNA  
; ORGANISM: Clavularia sp  
US-09-459-956-7

Query Match 40.0%; Score 271; DB 3; Length 801;  
Best Local Similarity 63.0%; Pred. No. 6.6e-85;  
Matches 418; Conservative 0; Mismatches 245; Indels 0; Gaps 0;

Qy 11 CCAAGAATGTTATCAAGGAGTTTCATGAGGTTTAAAGTTTCGCATGGAAGAAACGGTCAATG 70  
Db 125 CCATGGGTGTGATTAAACCAGACATGAAGATTAAAGCTGAAGATGAAGAAATGTAAACG 184  
Qy 71 GGCACGAGTTTGAATAGAAGGCGAAGGAGAGGGGAGGCCATACGAAGGCCCAATATACCG 130  
Db 185 GGCAATGTTTGTGATCGAAGGAGAGGAGAAAGCCCTTACGATGGGACACACACTT 244  
Qy 131 TAAAGCTTAAGGTAAACCAAGGGGGGACCTTTGCCATTTGCTTGGGATATTTTGTCAACAC 190  
Db 245 TAAACCTGGAAGTGAAGGAAGGTGGCCTCTGCCCTTTTCTTACGATATCTTGTCAACACG 304  
Qy 191 AATTTCAGTATGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCAGACTATAAAA 250  
Db 305 CGTTCCAGTACGGAACAGAGCATTGACAAAATACCCAGACGATATAGCAGACTATTTC 364  
Qy 251 AGCTGTCAATTTCTCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAAGACGGTGGCG 310  
Db 365 AGCAGTCGTTTCCGAGGGATATTTCTGGGAAAGAACCATGACTTTTGAAGACAAAGGCA 424  
Qy 311 TCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCACTCTACAAGGTCAAGT 370  
Db 425 TTGTCAAAGTGAAAAGTGACATAAGCATGGAGGAAGACTCCTTTATCTATGAAATTCGTT 484  
Qy 371 TCATTGGCGTGAACCTTTCTTCCGATGGACCTGTTATGCAAAAGAAAGACAATGGCTGGG 430  
Db 485 TTGATGGGATGAACCTTTCTCCCAATGGTCCGGTTATGCAGAAAAAAACTTTGAAGTGGG 544  
Qy 431 AAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAGATTCATAAGG 490  
Db 545 AACCATCCACTGAGATTATGTACGTGCGTGATGGAGTGTGTCGAGATATTAGCCATT 604  
Qy 491 CTCTGAAGCTGAAGACCGTGGTCACTTACCTAGTTGAATTCAAAAGTATTTACATGGCAA 550  
Db 605 CTCTGTTGCTGGAGGGAGGTGGCCATTACCGATGTGACTTCAAAAGTATTTACAAAGCAA 664  
Qy 551 AGAAGCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGATATAACAAGCC 610  
Db 665 AAAAAGTTGTCAAATTTGCCAGACTATCACTTTGTGGACCATCGCATTGAGATCTTGAACC 724  
Qy 611 ACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGAGCCACCATCTGT 670  
Db 725 ATGACAAGGATTACAACAAAGTAACGCTGTATGAGATGCAAGTTCAGTTGCTCGCTATCTTTC 784  
Qy 671 TCC 673  
Db 785 TGC 787

RESULT 5  
US-09-459-956-2  
; Sequence 2, Application US/09459956  
; Patent No. 6342379  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger Y.  
; APPLICANT: Gonzalez, Iii, Jesus E.  
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY  
; TITLE OF INVENTION: OPTICAL METHODS  
; FILE REFERENCE: REGEN1290-4  
; CURRENT APPLICATION NUMBER: US/09/459,956  
; CURRENT FILING DATE: 1999-12-13  
; PRIOR APPLICATION NUMBER: 08/765,860  
; PRIOR FILING DATE: 1999-05-08  
; PRIOR APPLICATION NUMBER: 08/481,977  
; PRIOR FILING DATE: 1995-06-07  
; PRIOR APPLICATION NUMBER: PCT/US96/09652  
; PRIOR FILING DATE: 1996-06-06  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 690  
; TYPE: DNA



```
; ORGANISM: Anemonia majano
US-09-459-956-2

Query Match      26.3%; Score 178.4; DB 3; Length 690;
Best Local Similarity 56.1%; Pred. No. 3.3e-52;
Matches 360; Conservative 0; Mismatches 276; Indels 6; Gaps 1;

QY 9 TTCCAAGAATGTTATCAAGGAGTTTCATGAGGTTTAAAGTTCGCATGGAAGAACGGTCAA 68
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9 TTCAAACAAGTTTATCGGAGATGACATGAAATGACCTACCATATGGATGGGTGTCAA 68

QY 69 TGGGCACGAGTTTGAAATAGAAGCGGAAGGAGAGGGGAGGCCATACGAAGGCCACAATAC 128
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 69 TGGGCATTACTTTACCGTCAAAGGTGAAGGCAACGGGAAGCCATACGAAGGCGCAGAC 128

QY 129 CGTAAAGCTTAAAGGTAA-----CCAAGGGGGGACCTTTGCCATTTCCTTGGGATATTT 182
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 129 TTCGACTTTTAAAGTCAACCATGGCCAAACGGTGGGCCCTTGCAATTCCTTTGACATACT 188

QY 183 GTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCAGA 242
   || || || || || || || || || || || || || || || || || || || || ||
Db 189 ATCTACAGTGTTCAAATATGGAATCGATGCTTTACTGCGTATCCTACCAGTATGCCCGA 248

QY 243 CTATAAAAAGCTGTCTTCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAAGA 302
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 249 CTATTTCAAACAAGCATTTCTCGACGGAATGTCAATATGAAGGACTTTTACCTATGAAGA 308

QY 303 CGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTACAA 362
   || || || || || || || || || || || || || || || || || || || || ||
Db 309 TGGAGGAGTTGCTACAGCCAGTTGGGAATAAGCCTTAAAGGCCAACTGCTTTGAGCACAA 368

QY 363 GGTCAAGTTTCATTGGCGTGAACTTTCCTTCCGATGGACCTGTTATGCAAAAGAAGACAAT 422
   || || || || || || || || || || || || || || || || || || || || ||
Db 369 ATCCACGTTTCATGGAGTGAACCTTCCTGCTGATGGACCTGTGATGGCGAAGAAGACAAC 428

QY 423 GGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAGAT 482
   || || || || || || || || || || || || || || || || || || || || ||
Db 429 TGGTTGGGACCCATCTTTTGAGAAAATGACTGTCTCGATGGAATATTGAAGGTTGATGT 488

QY 483 TCATAAGGCTCTGAAGCTGAAAGACGGTGGTGCATTACCTAGTTGAATTCAAAAGTATTTA 542
   || || || || || || || || || || || || || || || || || || || || ||
Db 489 CACCGCGTTCCTCATGCTGCAAGGAGGTGGCAATTACAGATGCCAATTCACACTTCTTA 548

QY 543 CATGGCAAAGAAGCCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGATAT 602
   || || || || || || || || || || || || || || || || || || || || ||
Db 549 CAAGACAAAAAAACCGGTGACGATGCCACCAACCATGTGGTGGAAACATCGCATTCGGAG 608

QY 603 AACAAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGA 644
   || || || || || || || || || || || || || || || || || || || || ||
Db 609 GACCGACCTTGACAAAAGGTGGCAACAGTGTTCAGCTGACGGA 650
```

```
RESULT 6
US-09-459-956-4
; Sequence 4, Application US/09459956
; Patent No. 6342379
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Gonzalez, III, Jesus E.
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY
; TITLE OF INVENTION: OPTICAL METHODS
; FILE REFERENCE: REGEN1290-4
; CURRENT APPLICATION NUMBER: US/09/459,956
; CURRENT FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: 08/765,860
; PRIOR FILING DATE: 1999-05-08
; PRIOR APPLICATION NUMBER: 08/481,977
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: PCT/US96/09652
; PRIOR FILING DATE: 1996-06-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
```

```
; LENGTH: 696
; TYPE: DNA
; ORGANISM: Zoanthus sp
US-09-459-956-4

Query Match      21.7%; Score 147.2; DB 3; Length 696;
Best Local Similarity 56.8%; Pred. No. 3.5e-41;
Matches 316; Conservative 0; Mismatches 228; Indels 12; Gaps 2;

QY 9 TTCCAAGAATGTTATCAAGGAGTTTCATGAGGTTTAAAGTTCGCATGGAAGGAACGGTCAA 68
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9 TTCAAAGCACGGTCTAAAAGAAGAAATGACAAATGAAATACCAATGGAAGGGTTCGTCAA 68

QY 69 TGGGCACGAGTTTGAAATAGAAGCGGAAGGAGAGGGGAGGCCATACGAAGGCCACAATAC 128
   || || || || || || || || || || || || || || || || || || || || ||
Db 69 CGGACATAAAATTTGTGATCACGGGCGAAGGCATTGGATATCCGTTCAAAGGGGAACAGAC 128

QY 129 CGTAAAGCTTAAAGTTAACCAAGGGGGGACCTTTTGCCATTTTGGGATATTTTGTCAAC 188
   || || || || || || || || || || || || || || || || || || || || ||
Db 129 TATTAAATCTGTGTGATCGAAGGGGGACCATTTGCCATTTTCCGAAGACATATTGTGAGC 188

QY 189 ACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCAGACTATAA 248
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 189 TGGCTTTAAGTACGGAGACAGGATTTTCACCTGAATATCCTCAAGACATAGTAGACTATTT 248

QY 249 AAAGCTGTCAATTTCTGAAGGATTTAAATGGGAAGGGTCAATGAACCTTTGAAGACGGTG- 307
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 249 CAAGAACTCGTGTCTCTGCTGGATATACATGGGCGAGGTCCTTTTCTCTTTGAGGATGGAGC 308

QY 308 -----GCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTACAA 362
   || || || || || || || || || || || || || || || || || || || || ||
Db 309 AGTCTGCATATGCAATGTAGATATAACAGTGAGTGTCAAAGAAAACCTGCATTTATCATAA 368

QY 363 GGTCAAGTTTCATTGGCGTGAACCTTTCCTTCCGATGGACCTGTATGCAAAAAGAAGACAAT 422
   || || || || || || || || || || || || || || || || || || || || ||
Db 369 GAGCATATTTAATGGAATGAATTTTCTGCTGATGGACCTGTGATGAAAAAAGATGACAAC 428

QY 423 GGGCTGGGAAGCCAGCACTGAGCGTTT-----GTATCCTCGTGATGGCGTGTGAAAGG 476
   || || || || || || || || || || || || || || || || || || || || ||
Db 429 TAACTGGGAAGCATCCTCGGAGAAGATCATGCCAGTACCTAAGCAGGGGATACTGAAAGG 488

QY 477 AGAGATTCATAAGGCTCTGAAGCTGAAAGCTGAAAGACGGTGGTGCATTACCTAGTTGAATTCAAAG 536
   || || || || || || || || || || || || || || || || || || || || ||
Db 489 GGATGTCTCCATGTACCTCCTTCTGAAGGATGGTGGGCGTTACCGGTGCCAGTTCGACAC 548

QY 537 TATTACATGGCAAAAG 552
   || || || || || || || || || || || || || || || || || || || || ||
Db 549 AGTTTACAAAGCAAAAG 564
```

```
RESULT 7
US-09-977-897-1
; Sequence 1, Application US/09977897
; Patent No. 6780974
; GENERAL INFORMATION:
; APPLICANT: Chen, Yih-Tai
; APPLICANT: Cao, Longguang
; TITLE OF INVENTION: A synthetic DNA encoding an orange seapen-derived green fluorescer
; TITLE OF INVENTION: protein with codon preference of mammalian expression systems and
; FILE REFERENCE: 41856-5
; CURRENT APPLICATION NUMBER: US/09/977,897
; CURRENT FILING DATE: 2001-10-15
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1482
; TYPE: DNA
; ORGANISM: Ptilosarcus gurneyi
US-09-977-897-1
```

```
Query Match      20.9%; Score 141.4; DB 4; Length 1482;
Best Local Similarity 51.3%; Pred. No. 6.5e-39;
Matches 328; Conservative 0; Mismatches 311; Indels 0; Gaps 0;
```

QY	13	AAGAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTTCGCATGGAAGGAACGGTCAATGGG	72
Db	37	AACACCGGCCTGAAGGAGATCATGAGCGCCAAGGCCAGCGTGGAGGSCATCGTGAACAAC	96
QY	73	CACGAGTTTGAAATAGAAGCGGAAGGAGAGGGGAGGCCATACGAAGSCACAATACCGTA	132
Db	97	CACGTGTTTCAGCATGGAGGGCTTCGGCAAGGGCAA	156
QY	133	AAGCTTAAGGTAAACCAAGGGGGACCTTTGCCATTTTGCTTGGGATATTTGTCAACCACAA	192
Db	157	CAGATCCGGGTGACCAAGGGCGGCCCTCTGCCCTTCGCCTTCGCACATCGTGAGCATCGCC	216
QY	193	TTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACAGACTATATAAAAG	252
Db	217	TTCCAGTACGGCAACCGGACCTTCACCAAGTATCCCGACGACATCGCCGACTACTTCTGTG	276
QY	253	CTGTCAATTTCCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAAGACGGTGGCGTC	312
Db	277	CAGAGCTTCCCTGCCGGCTTCTTCTACGAGCGGAACCTGCGGTTTCGAGACGGCGCCATC	336
QY	313	GTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTACAAGGTCAAGTTC	372
Db	337	GTGGACATCCGGAGCGACATCAGCCTGGAGGACGACAAAGTTCCTACTACAAGGTGGAGTAC	396
QY	373	ATTGGCGTGAACCTTTCCTCCGATGGACCTGTTATGCAAAAGAAAGACAATGGCTGGAA	432
Db	397	CGCGCAACGGCTTCCCTAGCAACGGCCCTGTGATGCAGAAGGCCCTCCTGGGCATGGAG	456
QY	433	GCCAGCACTGAGCGTTTGTATCCTCTGTATGCGCGTGTGAAAGGAGAGATTTCATAAGCT	492
Db	457	CCCAGCTTCGAGGTGGTGATCATGAACAGCGGCGTGCTGGTGGGCGAGGTGGACCTGGTG	516
QY	493	CTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATTTACATGGCAAAG	552
Db	517	TACAAGCTGGAGAGCGGCAACTATACAGCTGCCACATGAAGACCTTCTACCGGAGCAAG	576
QY	553	AAGCCTGTGCAGCTACAGGGTACTACTATGTTGACTCCAAACTGGATATAACAAGCCAC	612
Db	577	GGCGCGTGAAGGAGTTCCTGTAGTACCTATTTCATCCACCACCGGCTGGAGAAAGAACTAC	636
QY	613	AACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACC	651
Db	637	GTGGAGGAGGGCAGCTTCGTGGAGCAGCACGAGACCGGCC	675

RESULT 8  
US-09-459-956-3  
; Sequence 3, Application US/09459956  
; Patent No. 6342379  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger Y.  
; APPLICANT: Gonzalez, III, Jesus E.  
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY  
; TITLE OF INVENTION: OPTICAL METHODS  
; FILE REFERENCE: REGEN1290-4  
; CURRENT APPLICATION NUMBER: US/09/459,956  
; CURRENT FILING DATE: 1999-12-13  
; PRIOR APPLICATION NUMBER: 08/765,860  
; PRIOR FILING DATE: 1999-05-08  
; PRIOR APPLICATION NUMBER: 08/481,977  
; PRIOR FILING DATE: 1995-06-07  
; PRIOR APPLICATION NUMBER: PCT/US96/09652  
; PRIOR FILING DATE: 1996-06-06  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 696  
; TYPE: DNA  
; ORGANISM: Zoanthus sp  
US-09-459-956-3

Query Match 19.7%; Score 133.4; DB 3; Length 696;

[illegible]

## RESULT 9

```

US-09-609-161B-15
; Sequence 15, Application US/09609161B
; Patent No. 6436682
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; APPLICANT: Szent-Gyorgyi, Christopher
; APPLICANT: PROLUME, LTD.
; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PRO
; TITLE OF INVENTION: AND FLUORESCENT PROTEINS AN
; TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS
; FILE REFERENCE: 24729-121B
; CURRENT APPLICATION NUMBER: US/09/609,161B
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/277,716
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/102,939
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 60/089,367
; PRIOR FILING DATE: 1998-06-15
; PRIOR APPLICATION NUMBER: 60/079,624
; PRIOR FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Renilla mulleri
; FEATURE:

```

```
; NAME/KEY: CDS
; LOCATION: (259)..(975)
; OTHER INFORMATION: Renilla mulleri Green Fluorescent Protein (GFP)
US-09-609-161B-15

Query Match      19.6%; Score 133.2; DB 3; Length 1079;
Best Local Similarity 50.6%; Pred. No. 4.2e-36;
Matches 321; Conservative 0; Mismatches 313; Indels 0; Gaps 0;

QY 13 AAGAAATGTTATCAAGAGATTTCATGAGGTTTTAAGGTTCCGATGGAAGAACGGTCAATGGG 72
Db 280 AACACTTGTTTACAGAAGTAATGTCGTATAAAGTAAATCTGGAAGGAATTGTAAACAAC 339

QY 73 CACGAGTTTGAATAGAACGGCGAAGGAGAGGGGAGGCCATACGAAAGGCCACAATACCGTA 132
Db 340 CATGTTTTTACAATGGAGGGTTGCGGCAAAAGGGAATATTTATTCGGCAATCAACTGGTT 399

QY 133 AAGCTTAAGGTAACCAAGGGGGGACCTTTGCCATTTGCTTGGGATATTTTGTCAACCACAA 192
Db 400 CAGATTTCGTGTCACGAAAGGGCCCCCACTGCCCTTTGCATTTGATATTGTGCACCACT 459

QY 193 TTTTCAGTATGGAAGCRAAGGTATATGTCAAGCACCCCTGCCGACATACCAGACTATAAAAAG 252
Db 460 TTTCAATATGGCAACCGTACTTTTCACGAAATATCCGAATGATATATCAGATTATTTTATA 519

QY 253 CTGTCATTTCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAAGACGGTGGCGTC 312
Db 520 CAATCATTTCCAGCAGGATTTATGTATGAACGAACATTACGTTACGAAGATGGCGGACTT 579

QY 313 GTTACTGTAAACCCAGGATTCAGTTTTGCAGGATGGCTGTTTCATCTACAAGGTCAAAGTTC 372
Db 580 GTTGAAATTCGTTTCAGATATAAAATTTAATAGAACACAAGTTCGTCTACAGAGTGAATAC 639

QY 373 ATTGGCGTGAACCTTTCCCTCCGATGGACCTGTATGCAAAAAGACAAATGGGCTGGGAA 432
Db 640 AAAGGTAGTAACCTCCAGATGATGGTCCCGTCATGCAGAAGACTATCTTAGGAATAGAG 699

QY 433 GCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGGTGTTGAAAGGAGAGATTTCATAAGGCT 492
Db 700 CCTTCATTTGAAGCCATGTACATGAATAATGGCGTCTTGGTCGGCGAAGTAATCTTGTC 759

QY 493 CTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATTTACATGGCAAAG 552
Db 760 TATAAACTAAACCTCTGGGAAATATTATTTCATGTCACATGAAACACATTAAATGAAGTCGAAA 819

QY 553 AAGCCTGTGCAGCTACCAAGGCTACTACTATGTTGACTCCAAACTGGATATAACAAGCCAC 612
Db 820 GGTGTAGTAAAGGAGTTTCCTTCGTATCATTTTATTCAACATCGTTTGGAAAAGACTTAC 879

QY 613 AACGAAGACTATACAATCGTTGAGCAGTATGAAA 646
Db 880 GTAGAAGACGGGGGTTTCGTTGAACAGCATGAGA 913
```

```
RESULT 10
US-09-626-581D-64
; Sequence 64, Application US/09626581D
; Patent No. 6548249
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; TITLE OF INVENTION: Fusions of Scaffold Proteins with Random Peptide
; TITLE OF INVENTION: Libraries
; FILE REFERENCE: A-66900-3/RMS
; CURRENT APPLICATION NUMBER: US/09/626,581D
; CURRENT FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 09/169,015
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 09/415,765
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 64
; LENGTH: 1079
```

```
; TYPE: DNA
; ORGANISM: Renilla muelleri
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (259)..(975)
; OTHER INFORMATION:
US-09-626-581D-64

Query Match      19.6%; Score 133.2; DB 4; Length 1079;
Best Local Similarity 50.6%; Pred. No. 4.2e-36;
Matches 321; Conservative 0; Mismatches 313; Indels 0; Gaps 0;

QY 13 AAGAAATGTTATCAAGGAGTTTCATGAGGTTTTAAGGTTCCGATGGAAGGAACGGTCAATGGG 72
Db 280 AACACTTGTTTACAAGAAGTAATGTCGTATAAAGTAAATCTGGAAGGAATTGTAAACAAC 339

QY 73 CACGAGTTTGAATAGAACGGCGAAGGAGAGGGGAGGCCATACGAAAGGCCACAATACCGTA 132
Db 340 CATGTTTTTACAATGGAGGGTTGCGGCAAAAGGGAATATTTTATTCGGCAATCAACTGGTT 399

QY 133 AAGCTTAAGGTAACCAAGGGGGGACCTTTGCCATTTGCTTGGGATATTTTGTCAACCACAA 192
Db 400 CAGATTTCGTGTCACGAAAGGGCCCCCACTGCCCTTTGCATTTGATATTGTGCACCACT 459

QY 193 TTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCAGACTATAAAAAG 252
Db 460 TTTCAATATGGCAACCGTACTTTTCACGAAATATCCGAATGATATATCAGATTATTTTATA 519

QY 253 CTGTCATTTCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAAGACGGTGGCGTC 312
Db 520 CAATCATTTCCAGCAGGATTTATGTATGAACGAACATTACGTTACGAAGATGGCGGACTT 579

QY 313 GTTACTGTAAACCCAGGATTCAGTTTTGCAGGATGGCTGTTTCATCTACAAGGTCAAAGTTC 372
Db 580 GTTGAAATTCGTTTCAGATATAAAATTTAATAGAACACAAGTTCGTCTACAGAGTGAATAC 639

QY 373 ATTGGCGTGAACCTTTCCCTCCGATGGACCTGTATGCAAAAAGACAAATGGGCTGGGAA 432
Db 640 AAAGGTAGTAACCTCCAGATGATGGTCCCGTCATGCAGAAGACTATCTTAGGAATAGAG 699

QY 433 GCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGGTGTTGAAAGGAGAGATTTCATAAGGCT 492
Db 700 CCTTCATTTGAAGCCATGTACATGAATAATGGCGTCTTGGTCGGCGAAGTAATCTTGTC 759

QY 493 CTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATTTACATGGCAAAG 552
Db 760 TATAAACTAAACCTCTGGGAAATATTATTTCATGTCACATGAAACACATTAAATGAAGTCGAAA 819

QY 553 AAGCCTGTGCAGCTACCAAGGCTACTACTATGTTGACTCCAAACTGGATATAACAAGCCAC 612
Db 820 GGTGTAGTAAAGGAGTTTCCTTCGTATCATTTTATTCAACATCGTTTGGAAAAGACTTAC 879

QY 613 AACGAAGACTATACAATCGTTGAGCAGTATGAAA 646
Db 880 GTAGAAGACGGGGGTTTCGTTGAACAGCATGAGA 913
```

```
RESULT 11
US-09-415-765B-64
; Sequence 64, Application US/09415765B
; Patent No. 6548632
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; TITLE OF INVENTION: Fusions of Scaffold Proteins with Random Peptide
; TITLE OF INVENTION: Libraries
; FILE REFERENCE: A66900-1/RMS/AMS
; CURRENT APPLICATION NUMBER: US/09/415,765B
; CURRENT FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: 09/169,015
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 64
```



; LENGTH: 1079  
; TYPE: DNA  
; ORGANISM: Renilla muelleri  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (259)..(975)  
; OTHER INFORMATION:  
US-09-415-765B-64

Query Match 19.6%; Score 133.2; DB 4; Length 1079;  
Best Local Similarity 50.6%; Pred. No. 4.2e-36;  
Matches 321; Conservative 0; Mismatches 313; Indels 0; Gaps 0;

QY 13 AAGAATGTTATCAAGGAGTTTCATGAGGTTTAAAGTTTCGCATGGAAGGAACGGTCAATGGG 72  
Db 280 AACACTTGTTTACAGAAGTAATGTCGTATAAAGTAAATCTGGAAGGAATTGTAAACAAC 339  
QY 73 CACGAGTTTGAATAGAACGCGAAGGAGAGGGGAGGCCATACGAAGGCCACAATACCGTA 132  
Db 340 CATGTTTTTACATGAGGGTTGCGGCAAGGGAATATTTTATTCGGCAATCAACTGGTT 399  
QY 133 AAGCTTAAGGTAAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATTTTGTCAACCACAA 192  
Db 400 CAGATTTCGTGTCACGAAAGGGGCCCTTTCGCAATATCCGAATGATATATCAGATTATTTTATA 459  
QY 193 TTTCAATATGGCAACCGTACTTTTACGAAATATCCGAATGATATATCAGATTATTTTATA 519  
QY 253 CTGTCAATTCCTGAAGGATTAAATGGGAAAGGGTCAATGAACTTTGAAGACGGTGGCGTC 312  
Db 520 CAATCATTTCCAGCAGGATTATGATGAACGAACATTACGTTACGAAGATGGCGACTT 579  
QY 313 GTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTACAAGGTCAAGTTC 372  
Db 580 GTTGAAATTCGTTTCAGATATAAATTTAATAGAACAAGTTCGTTACAGAGTGGAAATAC 639  
QY 373 ATTGGCGTGAACCTTCCGATGGACCTGTTATGCAAAAGAAAGACAAATGGGCTGGAA 432  
Db 640 AAAGGTAGTAACTTCCAGATGATGGTCCCGTCATGCAGAAGACTATCTTAGGAATAGAG 699  
QY 433 GCCAGCCTGAGCGTTGTATCCTCGTATGGCGTGTGAAAGGAGAGATTTCATAAGGCT 492  
Db 700 CCTTCATTTGAAGCCATGTACATGAATAATGGCGTCTTGGTCGGCGAAGTAATCTTGTC 759  
QY 493 CTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATTTACATGGCAAAG 552  
Db 760 TATAAATAAACTCTGGGAAATATTTATTCATGTCACATGAAACATTTAATGAAGTCGAA 819  
QY 553 AAGCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGATATAACAAGCCAC 612  
Db 820 GGTGTAGTAAAGGAGTTTCCTTCGTATCATTTTATTCACATCGTTTGGAAAGACTTAC 879  
QY 613 AACGAAGACTATACAATCGTTGAGCAGTATGAAA 646  
Db 880 GTAGAAGACGGGGGTTTCGTTGAACAGCATGAGA 913

RESULT 12  
US-09-626-580C-64  
; Sequence 64, Application US/09626580C  
; Patent No. 6562617  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, David  
; APPLICANT: Bogenberger, Jakob M.  
; APPLICANT: Peele, Beau R.  
; TITLE OF INVENTION: FUSIONS OF SCAFFOLD PROTEINS WITH RANDOM PEPTIDE LIBRARIES  
; FILE REFERENCE: A-66900-2/RMS/AMS  
; CURRENT APPLICATION NUMBER: US/09/626,580C  
; CURRENT FILING DATE: 2000-07-27  
; PRIOR APPLICATION NUMBER: US 09/415,765  
; PRIOR FILING DATE: 1999-10-08  
; PRIOR APPLICATION NUMBER: US 09/169,015

; PRIOR FILING DATE: 1998-10-08  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 64  
; LENGTH: 1079  
; TYPE: DNA  
; ORGANISM: Renilla muelleri  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (259)..(975)  
; OTHER INFORMATION:  
US-09-626-580C-64

Query Match 19.6%; Score 133.2; DB 4; Length 1079;  
Best Local Similarity 50.6%; Pred. No. 4.2e-36;  
Matches 321; Conservative 0; Mismatches 313; Indels 0; Gaps 0;

QY 13 AAGAATGTTATCAAGGAGTTTCATGAGGTTTAAAGTTTCGCATGGAAGGAACGGTCAATGGG 72  
Db 280 AACACTTGTTTACAGAAGTAATGTCGTATAAAGTAAATCTGGAAGGAATTGTAAACAAC 339  
QY 73 CACGAGTTTGAATAGAACGCGAAGGAGAGGGGAGGCCATACGAAGGCCACAATACCGTA 132  
Db 340 CATGTTTTTACATGAGGGTTGCGGCAAGGGAATATTTTATTCGGCAATCAACTGGTT 399  
QY 133 AAGCTTAAGGTAAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATTTTGTCAACCACAA 192  
Db 400 CAGATTTCGTGTCACGAAAGGGGCCCTTTCGCAATATCCGAATGATATATCAGATTATTTTATA 459  
QY 193 TTTCAATATGGCAACCGTACTTTTACGAAATATCCGAATGATATATCAGATTATTTTATA 519  
QY 253 CTGTCAATTCCTGAAGGATTAAATGGGAAAGGGTCAATGAACTTTGAAGACGGTGGCGTC 312  
Db 520 CAATCATTTCCAGCAGGATTATGATGAACGAACATTACGTTACGAAGATGGCGACTT 579  
QY 313 GTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTACAAGGTCAAGTTC 372  
Db 580 GTTGAAATTCGTTTCAGATATAAATTTAATAGAACAAGTTCGTTACAGAGTGGAAATAC 639  
QY 373 ATTGGCGTGAACCTTCCGATGGACCTGTTATGCAAAAGAAAGACAAATGGGCTGGAA 432  
Db 640 AAAGGTAGTAACTTCCAGATGATGGTCCCGTCATGCAGAAGACTATCTTAGGAATAGAG 699  
QY 433 GCCAGCCTGAGCGTTGTATCCTCGTATGGCGTGTGAAAGGAGAGATTTCATAAGGCT 492  
Db 700 CCTTCATTTGAAGCCATGTACATGAATAATGGCGTCTTGGTCGGCGAAGTAATCTTGTC 759  
QY 493 CTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATTTACATGGCAAAG 552  
Db 760 TATAAATAAACTCTGGGAAATATTTATTCATGTCACATGAAACATTTAATGAAGTCGAA 819  
QY 553 AAGCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGATATAACAAGCCAC 612  
Db 820 GGTGTAGTAAAGGAGTTTCCTTCGTATCATTTTATTCACATCGTTTGGAAAGACTTAC 879  
QY 613 AACGAAGACTATACAATCGTTGAGCAGTATGAAA 646  
Db 880 GTAGAAGACGGGGGTTTCGTTGAACAGCATGAGA 913

RESULT 13  
US-09-277-716-15  
; Sequence 15, Application US/09277716A  
; Patent No. 6232107  
; GENERAL INFORMATION:  
; APPLICANT: Bryan, Bruce  
; APPLICANT: Szent-Gyorgyi, Christopher  
; APPLICANT: PROLUME, LTD.  
; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE  
; CURRENT APPLICATION NUMBER: US/09/277,716A  
; CURRENT FILING DATE: 1999-03-26



```
; EARLIER APPLICATION NUMBER: 60/102,939
; EARLIER FILING DATE: 1998-10-01
; EARLIER APPLICATION NUMBER: 60/089,367
; EARLIER FILING DATE: 1998-06-15
; EARLIER APPLICATION NUMBER: 60/079,624
; EARLIER FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Renilla mulleri
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (259)..(975)
; FEATURE:
; OTHER INFORMATION: Renilla mulleri Green Fluorescent Protein (GFP)
US-09-277-716-15

Query Match          19.6%; Score 133.2; DB 3; Length 1085;
Best Local Similarity 50.6%; Pred. No. 4.2e-36;
Matches 321; Conservative 0; Mismatches 313; Indels 0; Gaps 0;

QY 13 AAGAAATGTTATCAAGGAGTTCATGAGGTTTAAAGTTTCGCATGGAAGGACGGTCAATGGG 72
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 280 AACACTTGTTTACAGAAAGTAATGTCGTATAAAGTAAATCTGGAAGGAATTGTAAACAAC 339

QY 73 CACGAGTTTGAATAGAAAGCGGAAGGAGAGGGGAGGCCATACGAAAGGCCACAATACCGTA 132
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 340 CATGTTTTTACAATGGAGGGTTGCGGCAAAAGGGAATATTTTATTCGGCAATCAACTGGTT 399

QY 133 AAGCTTAAAGGTAAACCAAGGGGGGACCTTTGCCATTGCTTGGGATATTTTGTCAACCACAA 192
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 400 CAGATTCTGTCTACGAAAGGGGCCCCACTGCCCTTTGCATTTGATATTGTGTCAACCGCT 459

QY 193 TTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCAGACTATAAAAG 252
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 460 TTTCAATATGGCAACCGTACTTTTCACGAAATATCCGAATGATATATCAGATTATTTTATA 519

QY 253 CTGTCATTTCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTTGAAGACGGTGGCGTC 312
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 520 CAATCATTTCCAGCAGGATTTATGTATGAACGCAACATTACGTTACGAAGATGGCGGACTT 579

QY 313 GTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTTCATCTACAAGTCAAGTTC 372
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 580 GTTGAAATTCGTTTCAGATATAAAATTTTAATAGAAACAAGTTTCGTCTACAGAGTGAATAC 639

QY 373 ATTGGCGTGAACTTTTCCTTCCGATGGACCTGTGTATGCAAAAAGACAAATGGGCTGGGAA 432
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 640 AAAGGTAGTAACCTTCCAGATGATGGTCCCGTCATGCAGAAAGACTATCTTAGGAATAGAG 699

QY 433 GCCAGCACTGAGCGTTTGATCCTCGTGATGGCGTGTGAAAGGAGAGATTCATAAGGCT 492
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 700 CCTTCATTTGAAGCCCATGTACATGAATAATGGCGTCTTGGTCGGCGGAAGTAATTCTTGTC 759

QY 493 CTGAAGCTGAAAGACGCGTGGTCATTACCTAGTTGAATTCAAAAGTATTTTACATGGCAAAG 552
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 760 TATAAACTAAACTCTGGGAATATTTATTCATGTACATGAAACATTAAATGAAGTCGAAA 819

QY 553 AAGCCTGTGCAGCTACACAGGGTACTACTATGTTGACTCCAAACTGGATATAACAAGCCAC 612
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 820 GGTGTAGTAAAGGAGTTTCCTTCGTATCATTTTATTCAACATCGTTTGGAAAGACTTAC 879

QY 613 AACGAAGACTATACAAATCGTTGAGCAGTATGAAA 646
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 880 GTAGAAGACGGGGGTTTCGTTGAACAGCATGAGA 913
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 14
US-09-277-716-30
; Sequence 30, Application US/09277716A
; Patent No. 6232107
; GENERAL INFORMATION:
```

```
; APPLICANT: Bryan, Bruce
; APPLICANT: Szent-Gyorgyi, Christopher
; APPLICANT: PROLUME, LTD.
; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE L
; CURRENT APPLICATION NUMBER: US/09/277,716A
; CURRENT FILING DATE: 1999-03-26
; EARLIER APPLICATION NUMBER: 60/102,939
; EARLIER FILING DATE: 1998-10-01
; EARLIER APPLICATION NUMBER: 60/089,367
; EARLIER FILING DATE: 1998-06-15
; EARLIER APPLICATION NUMBER: 60/079,624
; EARLIER FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 1104
; TYPE: DNA
; ORGANISM: Ptilosarcus gurneyi
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (34)..(747)
; FEATURE:
; OTHER INFORMATION: Ptilosarcus Green Fluorescent Protein (GFP) (insert A)
US-09-277-716-30
```

```
Query Match          19.5%; Score 132.4; DB 3; Length 1104;
Best Local Similarity 50.3%; Pred. No. 8.1e-36;
Matches 325; Conservative 0; Mismatches 321; Indels 0; Gaps 0;

QY 13 AAGAATGTTATCAAGGAGTTCATGAGGTTTAAAGTTTCGCATGGAAGGAACGGTCAATGGG 72
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 55 AACACTGGACTGAAAGAGATTATGTCGGCAAAAGCTAGCGTTGAAGGAATCGTGAACAAT 114

QY 73 CACGAGTTTGAATAAGAGGCGAAGGAGAGGGGAGGCCATACGAAAGGCCACAATACCGTA 132
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 115 CACGTTTTTCCATGGAAGGATTTGGAAAAAGGCAATGTATTATTGGAAAAACCAATTGATG 174

QY 133 AAGCTTAAAGTAAACCAAGGGGGGACCTTTGCCATTTGCTTGGGATATTTTGTCAACCACAA 192
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175 CAAATCCGGGTTACAAAGGGAGGTCGGTTGCCATTCGCTTTCGATATTTGTTCCATAGCT 234

QY 193 TTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCAGACTATAAAAAAG 252
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 235 TTCCAATACGGGAATCGCACTTTTCACGAAATACCCAGACGACATTCGGGACTACTTTGTT 294

QY 253 CTGTCAATTCCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAAGACGGTGGCGTC 312
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 295 CAATCATTTCCCGGCTGGATTTTCTACGAAAGAAATCTACGCTTTGAAGATGGCGCCATT 354

QY 313 GTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTTCATCTACAAGGTCAAGTTC 372
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 355 GTTGACATTCGTTTCAGATATAAGTTTAGAAGATGATAAGTTCCTACTACAAAAGTGGAGTAT 414

QY 373 ATTGGCGTGAACCTTTCCCTTCCGATGGACCTGTTTATGCAAAAAGACAAATGGGCTGGGAA 432
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 415 AGAGGCAACGGTTTCCCTAGTAACGAGACCCGTGATGCAAAAAGCCCATCCTCGGCATGGAG 474

QY 433 GCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAGAGATTTCATAAGGCT 492
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 475 CCATCGTTTGAGGTGGTCTACATGAACAGCGGCGGCTTCTGGTGGGCGAAGTAGATCTCGTT 534

QY 493 CTGAAGCTGAAAGACGCGTGGTCAATTACCTAGTTGAATTCAAAAGTATTTTACATGGCAAAG 552
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 535 TACAAACTCGAGTCAGGGAACCTATTACTCGTGCCACATGAAACCGTTTTACAGATCCAAA 594

QY 553 AAGCCTGTGCAGCTACACAGGGTACTACTATGTTGACTTCCAACTGGATATAACAAGCCAC 612
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 595 GGTGGAGTGAAGAAGAAATCCCGGAATATCACTTTATCCATCATCTGCTGTGGAGAAAACCTAC 654

QY 613 AACGAAGACTATACAATCGTTGAGCAGTATGAAAAGAACCGAGGGAC 658
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 655 GTGGAAGAGGAAGCTTCGTGGAAACAACACGAGACGCGGCCATTGCGAC 700
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 15  
US-09-609-161B-30  
; Sequence 30, Application US/09609161B  
; Patent No. 6436682  
; GENERAL INFORMATION:  
; APPLICANT: Bryan, Bruce  
; APPLICANT: Szent-Gyorgyi, Christopher  
; APPLICANT: PROLUME, LTD.  
; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE LUC  
; TITLE OF INVENTION: AND FLUORESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIG  
; TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS  
; FILE REFERENCE: 24729-121B  
; CURRENT APPLICATION NUMBER: US/09/609,161B  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: 09/277,716  
; PRIOR FILING DATE: 1999-03-26  
; PRIOR APPLICATION NUMBER: 60/102,939  
; PRIOR FILING DATE: 1998-10-01  
; PRIOR APPLICATION NUMBER: 60/089,367  
; PRIOR FILING DATE: 1998-06-15  
; PRIOR APPLICATION NUMBER: 60/079,624  
; PRIOR FILING DATE: 1998-03-27  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn ver. 2.0  
; SEQ ID NO 30  
; LENGTH: 1104  
; TYPE: DNA  
; ORGANISM: Ptilosarcus gurneyi  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (34)..(747)  
; OTHER INFORMATION: Ptilosarcus Green Fluorescent Protein (GFP) (insert A)  
US-09-609-161B-30

Query Match 19.5%; Score 132.4; DB 3; Length 1104;  
Best Local Similarity 50.3%; Pred. No. 8.1e-36;  
Matches 325; Conservative 0; Mismatches 321; Indels 0; Gaps 0;

QY	13	AAGAATGTTATCAAGGAGTTTCATGAGGTTTAAAGTTCGCATGGAAGGAACGGTCAATGGG	72
Db	55	AACACTGGACTGAAAGAGATTATGTCGGCAAAAGCTAGCGTTGAAGGAATCGTGAACAAT	114
QY	73	CACGAGTTTGAATAGAAAGCGAAGGAGAGGGAGGCCCATACGAAGGCCACAATACCGTA	132
Db	115	CACGTTTTTCCATGGAAGGATTGGAAGGCAATGTATTATTTGGAACCAATGTATG	174
QY	133	AAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTGCTTGGGATATTTGTCAACCACAA	192
Db	175	CAATCCGGGTTACAAAGGGAGGTCGGTTGCCATTGCTTTTCGATATTGTTCCATAGCT	234
QY	193	TTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCTGCCGACATACCAGACTATAAAAAG	252
Db	235	TTCCAATACGGGAATCGCACTTTCACGAAATACCCAGACGACATTGCGGACTACTTTGTT	294
QY	253	CTGTCAATTTCCGTAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAAGACGGTGGCGTC	312
Db	295	CAATCATTTCCCGCTGGATTTTCTACGAAGAATACTACGCTTTGAAGATGGCGCCATT	354
QY	313	GTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTACAGGTCAAGTTC	372
Db	355	GTGACATTGCTTCAGATATAAGTTTAGAAGATGATAAGTTCCACTACAAAGTGGAGTAT	414
QY	373	ATTGGCGTGAACCTTTCCCTCCGATGGACCTGTTATGCAAAAGACATGGGCTGGAA	432
Db	415	AGAGGCAACGGTTTCCCTAGTAGAACGGACCCCGTATGCAAAAGCCATCTCCGCGATGGAG	474
QY	433	GCCAGCACTGAGCGTTTGTATCCTCGTATGGCGTGTGAAAGGAGAGATTTCATAAGGCT	492
Db	475	CCATCGTTTGAGGTGCTACATGAACAGCGCGCTTCTGTTGGCGGAAGTAGATCTCGTT	534
QY	493	CTGAAGCTGAAAGACGGTGGTGCATTACCTAGTTGAATTCAAAAGTATTTACATGGCAAAG	552

Db	535	TACAAACTCGAGTCAGGGAACCTATTACTCGTGCCACATGAAAAACGTTTACAGATCCAAA	594
QY	553	AAGCCTGTGCAGCTACCAAGGGTACTACTATGTTGACTCCAAACTGGATATAACAAGCCAC	612
Db	595	GGTGGAGTGAAGAATTCCTCGGAATATCACTTTATCCATCATCGTCTGGAGAAAAACCTAC	654
QY	613	AACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCAGGGGAC	658
Db	655	GTGGAAGGAAGGCTTCGTGGAAACAACACGAGACGGGCCATTGCGAC	700

Search completed: July 1, 2005, 18:23:04  
Job time : 173 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: June 30, 2005, 15:50:30 ; Search time 37.5 Seconds  
(without alignments)  
3479.198 Million cell updates/sec

Title: US-10-006-922A-11  
Perfect score: 1215  
Sequence: 1 atgagggtcttccaagaatgt.....gccaccatctgttcctttaa 678

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US10006922/runat\_30062005\_110552\_9435/app\_query.fasta\_1.839  
-DB=PIR -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10006922 @CGN\_1\_1\_63 @runat\_30062005\_110552\_9435 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR\_79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	213.5	17.6	238	1 JQ1514	green-fluorescent
2	93.5	7.7	26926	1 I38344	titin, cardiac mus
3	92	7.6	1873	2 T30944	surface protein pr
4	91	7.5	458	1 A42386	hes 90-binding pro
5	90	7.4	568	2 T06489	probable peptidylp
6	89	7.3	1433	1 A36734	bacillopeptidase F
7	87.5	7.2	1116	2 B70476	hypothetical prote
8	87	7.3	187	2 T01852	probable blue copp
9	87	7.2	1484	2 C97196	probable membrane
10	85.5	7.0	456	2 G69397	signal-transducing
11	85.5	7.2	1206	2 E86445	hypothetical prote
12	85	7.0	271	2 F69442	hypothetical prote
13	84.5	7.0	679	2 A40351	adhesion-type prot
14	84.5	7.0	725	1 IJMSNG	neural cell adhesi

15	84.5	7.0	2004	2 AC0314	probable membrane
16	84	7.0	839	2 T04859	extensin homolog F
17	84	6.9	1616	2 T17884	S-layer protein -
18	83.5	6.9	268	2 E90276	conserved hypothet
19	83.5	6.9	340	2 E69544	hypothetical prote
20	83.5	6.9	374	2 T06245	gibberellin 3 beta
21	83.5	6.9	374	2 T06244	gibberellin 3 beta
22	83	6.8	373	2 T50605	hypothetical prote
23	83	6.8	602	2 E81439	threonine-tRNA lig
24	83	6.8	15281	2 S41309	cyclosporin synthe
25	82.5	6.8	559	1 S55383	peptidylprolyl iso
26	82.5	6.9	589	2 E71251	probable oligoendo
27	82.5	6.8	680	2 S17982	Kallmann syndrome
28	82.5	6.8	751	1 I49497	transforming prote
29	82.5	6.8	931	2 T32919	hypothetical prote
30	82	6.7	341	2 B53125	restriction enzyme
31	81.5	6.7	292	2 C69106	glucose-1-phosphat
32	81.5	6.7	296	2 I37989	La 4.1 protein - h
33	81.5	6.7	551	1 S72485	peptidylprolyl iso
34	81.5	6.7	629	2 C64180	hypothetical prote
35	81.5	6.7	6805	2 S20901	titin - rabbit (fr
36	81	6.7	346	2 S77025	nitrilase (EC 3.5.
37	81	6.7	399	2 T40831	probable guanine n
38	81	6.7	587	1 E69171	phosphoesterase-re
39	80.5	6.6	404	1 S03849	ribonucleoprotein
40	80.5	6.6	415	1 JC1494	ribonucleoprotein
41	80.5	6.7	450	1 MHDG	Ig mu chain C regi
42	80.5	6.6	862	2 F75116	hypothetical prote
43	80.5	6.6	1003	1 AJCHPR	phosphoribosylamin
44	80.5	6.6	1115	1 IJMSNL	neural cell adhesi
45	80	6.6	621	2 A95250	choline binding pr

ALIGNMENTS

RESULT 1

JQ1514  
green-fluorescent protein [validated] - hydromedusa (Aequorea victoria)  
C;Species: Aequorea victoria  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C;Accession: JS0692; JQ1514; PQ0335; S48693; S51330; S51331  
R;Prasher, D.C.; Eckenrode, V.K.; Ward, W.W.; Prendergast, F.G.; Cormier, M.J.  
Gene 111, 229-233, 1992  
A;Title: Primary structure of the Aequorea victoria green-fluorescent protein.  
A;Reference number: JQ1514; MUID:92175527; PMID:1347277  
A;Accession: JS0692  
A;Molecule type: DNA  
A;Residues: 1-107,'S',109-238 <PRA1>  
A;Cross-references: UNIPROT:P42212; UNIPROT:Q17106; UNIPROT:Q17105; GB:M62654; NID:g1556  
A;Accession: JQ1514  
A;Molecule type: mRNA  
A;Residues: 1-99,'F',101-140,'L',142-218,'V',220-238 <PRA2>  
A;Cross-references: GB:M62653; NID:g155660; PIDN:AAA27721.1; PID:g155661  
A;Accession: PQ0335  
A;Molecule type: protein  
A;Residues: 46-64;74-122;132-151;154-183;185-200 <PRA3>  
R;Inouye, S.; Tsuji, F.I.  
FEBS Lett. 351, 211-214, 1994  
A;Title: Evidence for redox forms of the Aequorea green fluorescent protein.  
A;Reference number: S48693; MUID:94364470; PMID:8082767  
A;Accession: S48693  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-24,'Q',26-156,'P',158-171,'K',173-238 <INO>  
A;Cross-references: GB:L29345; NID:g606383; PIDN:AAA58246.1; PID:g606384  
R;Watkins, J.N.; Campbell, A.K.  
submitted to the EMBL Data Library, January 1995  
A;Reference number: S51330  
A;Accession: S51330  
A;Molecule type: mRNA  
A;Residues: 1-13,'V',15-24,'Q',26-44,'N',46-153,'G',155-156,'P',158-171,'K',173-227,'R',  
A;Cross-references: EMBL:X83959; NID:g634008; PIDN:CAA58789.1; PID:g634009  
A;Experimental source: clone gfp1





submitted to the Brookhaven Protein Data Bank, February 1996  
A;Reference number: A66736; PDB:1TIT  
A;Contents: annotation; conformation by (1)H-NMR, residues 5253-5341  
R;Pfuhl, M.; Pastore, A.  
submitted to the Brookhaven Protein Data Bank, August 1996  
A;Reference number: A66201; PDB:1NCT  
A;Contents: annotation; conformation by (1)H-NMR, residues 'S', 26059-26155  
C;Genetics:  
A;Gene: GDB:TTN  
A;Cross-references: GDB:127867; OMIM:188840  
A;Map position: 2q31-2q32  
C;Function:  
A;Description: structural protein forming filaments in striated muscle  
C;Superfamily: titin; fibronectin type III repeat homology; immunoglobulin homology; pro  
C;Keywords: alternative splicing; calmodulin binding; cardiac muscle; duplication; glyco  
structural protein  
F;24752-25008/Domain: protein kinase homology <KIN>  
F;84,177,905,2276,2378,2459,2481,2563,2669,2763,2896,3088,3179,3384,3432,3628,3772,4068,  
98,11066,11488,11515,11635,11949,12170,12478,12526,12645,12875,13001,13036,13295,13540,1  
tatus predicted  
F;16780,16976,17579,17602,17667,17681,17845,17899,18121,18188,18209,18336,18670,18680,18  
,21900,21935,22295,22495,22627,22897,23024,23318,23883,24012,24177,24290,24447,24642,248  
F;26171,26178,26184,26190/Binding site: phosphate (Ser) (covalent) #status experimental

Alignment Scores:  
Pred. No.: 2.81 Length: 26926  
Score: 93.50 Matches: 51  
Percent Similarity: 34.55% Conservative: 34  
Best Local Similarity: 20.73% Mismatches: 84  
Query Match: 7.70% Indels: 77  
DB: 1 Gaps: 13

US-10-006-922A-11 (1-678) x I38344 (1-26926)

QY 43 AAGGTTTCGATGGAAGAACGGTCAATGGGCACGAGTTTGAA-----ATA 87  
Db 23493 ArgTyrArgSerThrGlyLeuThrGluGlyLeuGlyTyrGluHisArgValThrAlaIle 23512

QY 88 GAAGCGAAGGAGAGGGAGGCCATACGAAGGCCACAATACCGTA----- 132  
Db 23513 AsnAlaArgGlySerGlyLysProSerArgProSerLysProIleValAlaMetAspPro 23532

QY 133 -----AAGCTTAAGGTAACCAAGGGGGACCTTTG 162  
Db 23533 IleAlaProProGlyLysProGlnAsnProArgValThrAspThrThrArgThrSerVal 23552

QY 163 CCATTGCTTGGATATTTTGTCAACCACAATTCAGTATGGAAGCAAGTA-----TAT 216  
Db 23553 SerLeuAlaTrpSerVal-----ProGluAspGluGlyGlySerLysValThrGlyTyr 23570

QY 217 GTC-----AAGCAC-----CCTGCC 231  
Db 23571 LeuIleGluMetGlnLysValAspGlnHisGluTrpThrLysCysAsnThrThrProThr 23590

QY 232 GACATACCAGACTATAAAAGCTGTCATTTCTCGAAGGATTAAATGGGAA---AGGGTC 288  
Db 23591 LysIleArgGluTyrThrLeuThrHisLeuProGlnGlyAlaGluTyrArgPheArgVal 23610

QY 289 ATGAACTTTGAACCGGT-----GGCGTCGTTACTGTA 321  
Db 23611 LeuAlaCysAsnAlaGlyGlyProGlyGluProAlaGluValProGlyThrValLysVal 23630

QY 322 ACCCAG-----GATTCAGTTTGCAGATGGCTGTTTC 354  
Db 23631 ThrGluMetLeuGluTyrProAspTyrGluLeuAspGluArgTyrGlnGluGlyIlePhe 23650

QY 355 ATCTACAAGGTCAGTTCAATTGGCGTGAACTTTCTCCGATGGA-----CCTGTT 405  
Db 23651 ValArgGlnGlyGlyValIleArgLeuThrIleProIleLysGlyLysProPheProIle 23670

QY 406 ATGCAAAAGAACAAATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGCC 465  
Db 23671 CysLys-----TrpThrLysGluGlyGlnAspIleSerLysArgAlaMet 23685

QY 466 GTGTTGAAGGAGAGATTTCATAAGGCTCTGAAGCTGAAA-----GACGGT 510  
Db 23686 IleAlaThrSerGluThrHisThrGluLeuValIleLysGluAlaAspArgGlyAspSer 23705

QY 511 GGTCAATTACCTAGTTGAATTC-----AAAAGTATTTCATGCGCAAAG 552  
Db 23706 GlyThrTyrAspLeuValLeuGluAsnLysCysGlyLysLysAlaValTyrIleLysVal 23725

QY 553 AAGCCTGTGCAGCTACCA 570  
Db 23726 ArgValIleGlySerPro 23731

RESULT 3  
T30944  
surface protein precursor - Enterococcus faecalis  
C;Species: Enterococcus faecalis  
C;Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T30944  
R;Shankar, V.; Baghdayan, A.S.; Huycke, M.M.; Lindahl, G.; Gilmore, M.S.  
Infect. Immun. 67, 193-200, 1999  
A;Title: Infection-derived Enterococcus faecalis strains are enriched in esp, a gene enc.  
A;Reference number: Z20943; MUID:99081742; PMID:9864215  
A;Accession: T30944  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-1873 <SHA>  
A;Cross-references: UNIPROT:Q9Z4N7; EMBL:AF034779; NID:G3873186; PID:G3873187; PIDN:AADO

Alignment Scores:  
Pred. No.: 2.17 Length: 1873  
Score: 92.00 Matches: 46  
Percent Similarity: 35.75% Conservative: 33  
Best Local Similarity: 20.81% Mismatches: 56  
Query Match: 7.57% Indels: 86  
DB: 2 Gaps: 13

US-10-006-922A-11 (1-678) x T30944 (1-1873)

QY 88 GAAGGCGAAGGAGAGGGAGGCCATACGAAGGCCACAATACCGTAAAGTTAAGGTAACC 147  
Db 708 LysGlyValGlyGluSerGluProIleThrGly-----ThrIleProIleLysIleVal 725

QY 148 AAGGGGGACCTTTGCCATTTGCTTGGGATATTTTGTCAACCACAATTTTCAGTATGGAAGC 207  
Db 726 -----TyrGlnAspGlySer 730

QY 208 -----AAGGTATATGTCAAGCACCTGCC 231  
Db 731 ValGlyThrThrAspLeuAlaValThrValSerLysAsnIleTyrGluAsnProGlyGlu 750

QY 232 GACATACCA---GACTATAAAAGCTGTCATTTCTGAAGGATTTAATGGGAAAGGGTC 288  
Db 751 AsnIleProAlaGlyTyrHisLysValThrPheThrAlaGly----- 764

QY 289 ATGAACTTTGAAGACCGTGGCGTCGTTACTGTAAACCAGGATTCAGTTTCAGGATGGC 348  
Db 765 -----GluGlyThrSerIleGluSerGly 772

QY 349 TGTTTCATCTACAAGGTCAAGTTTCATTGGCGTGAACCTTTCTTCGATGGA---CCTGTT 405  
Db 773 ThrThrValPheAlaValLys---AspGlyValSerLeuProGluAspLysLeuProVal 791

QY 406 ATGCAAAAGAACAAATGGGCTGGGAAGCCAGCACACTGAGCGTTTGTATCCTCGTGATGGC 465  
Db 792 LeuLysAlaLys-----AspGlyTyrThrAspAlaLysTrpProGluGluAla 807

QY 466 GTG-----TTGAAAGGAGAG-----ATTCAATAGGCTCTGAAGCTGAAAGAC 507  
Db 808 ThrGlnProIleLysAlaAspAspThrGluPheValSerSerAlaThrLysLeuAspAsp 827

QY 508 -----GGTGGTCATTACCTAGTTGAATTC----- 531

Db 828 IleIleGluAsnProGlyAspAsnIleProAlaGlyTyrHisLysValThrPheThrAla 847

QY 532 -----AAAAGTATTTACATGGCAAGAAGCCTGTGCAG 564

Db 848 GlyGluGlyThrSerIleGluSerGlyThrThrValPheAlaValLysAspGlyValSer 867

QY 565 CTACCAAGGTACTACTATGTTGACTCCAACTGGATATAACAAGCCACAACGAAGACTAT 624

Db 868 LeuPro-----GluAspLysLeuProValLeuLysAlaLysAspGlyTyr 882

QY 625 ACA 627

Db 883 Thr 883

RESULT 4

A42386

hsp 90-binding protein p59 - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004

C:Accession: A42386

R:Lebeau, M.C.; Massol, N.; Herrick, J.; Faber, L.E.; Renoir, J.M.; Radanyi, C.; Baulieu J. Biol. Chem. 267, 4281-4284, 1992

A:Title: P59, an hsp 90-binding protein. Cloning and sequencing of its cDNA and preparat

A:Reference number: A42386; MUID:92165768; PMID:1537818

A:Accession: A42386

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-458 <LEB>

A:Cross-references: UNIPROT:P27124; GB:M84988; NID:g165594; PIDN:AAA31439.1; PID:g165595

A:Experimental source: liver

A>Note: sequence extracted from NCBI backbone (NCBIP:83839)

C:Superfamily: human FK506-binding protein FKBP51; BKBP-type peptidylprolyl isomerase ho

F:50-97/Domain: BKBP-type peptidylprolyl isomerase homology <TT1>

F:319-352/Domain: tetratricopeptide repeat homology <TT1>

F:353-386/Domain: tetratricopeptide repeat homology <TT2>

Alignment Scores:

Pred. No.:	1.98	Length:	458
Score:	91.00	Matches:	42
Percent Similarity:	39.77%	Conservative:	26
Best Local Similarity:	24.56%	Mismatches:	49
Query Match:	7.49%	Indels:	54
DB:	1	Gaps:	10

US-10-006-922A-11 (1-678) x A42386 (1-458)

QY 169 GCTTGGGATATT-----TTGTCA 186

Db 89 AlaTrpAspIleAlaValAlaThrMetLysValGlyGluLeuCysArgIleThrCysLys 108

QY 187 CCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCAGACTAT 246

Db 109 ProGluTyrAlaTyrGlySer-----AlaGlySerProProLysIleProProAsn 125

QY 247 AAAAAGCTGTCAATTTCTCTGAA-----GGATTTAAATGGGAAAGGTCATGAACCTTT 297

Db 126 AlaThrLeuValPheGluValGluLeuPheGluPheLysGlyGluAspLeuThrAspAsp 145

QY 298 GAAGACGGTGGCGCTTACTGTAAACCAGGATCCAGTTTTCAGGATGGCTGTTTCATC 357

Db 146 GluAspGlyGlyIle-----Ile 151

QY 358 TACAAGGTCAAGTTCATTTGGCGTGAACCTTT-----CCTTCGATGGACCTGTTATGCAA 411

Db 152 ArgArgIleArgThrArgGlyGluGlyTyrAlaArgProAsnAspGlyAlaIleValGlu 171

QY 412 AAGAAGACAATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCT-----CGT 459

Db 172 ValAlaLeuGluGlyTyr-----TyrLysAspArgLeuPheAspGlnArgGluLeuArg 189

QY 460 GATGGCGTGTGAAAGGAGAG-----ATTCAAGGCTCTG--- 495

Db 190 PheGluValGlyGluGlyGluSerLeuAspLeuProCysGlyLeuGluLysAlaIleGln 209

QY 496 AAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAGTATTTTAC----- 543

Db 210 ArgMetGluLysGlyGluHisSerIleLeuTyrLeuLysProSerTyrAlaPheGlyAsn 229

QY 544 ATGGCAAGAAGCCTGTGCAGCTACCGAGGGTAC 576

Db 230 AlaGlyLysGluLysPheGlnIleProProTyr 240

RESULT 5

T06489

probable peptidylprolyl isomerase (EC 5.2.1.8) FKBP77 - wheat

C:Species: Triticum aestivum (common wheat)

C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004

C:Accession: T06489

R:Brieman, A.

submitted to the EMBL Data Library, August 1996

A:Reference number: Z15713

A:Accession: T06489

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-568 <BRI>

A:Cross-references: UNIPROT:O04843; EMBL:Y07636; PIDN:CAA68913.1

A:Experimental source: cv. ATIR, 2 day old plants, root tips

C:Genetics:

A:Gene: FKBP77

C:Superfamily: peptidylprolyl isomerase ROF1; BKBP-type peptidylprolyl isomerase homology,

C:Keywords: calmodulin binding; cis-trans-isomerase; cyclosporin A binding

F:54-101/Domain: BKBP-type peptidylprolyl isomerase homology <PPI>

F:169-212/Domain: BKBP-type peptidylprolyl isomerase homology <PPI2>

F:286-334/Domain: BKBP-type peptidylprolyl isomerase homology <PPI1>

F:482-515/Domain: tetratricopeptide repeat homology <PTR>

Alignment Scores:

Pred. No.:	2.59	Length:	568
Score:	90.00	Matches:	63
Percent Similarity:	33.22%	Conservative:	33
Best Local Similarity:	21.80%	Mismatches:	84
Query Match:	7.41%	Indels:	109
DB:	2	Gaps:	15

US-10-006-922A-11 (1-678) x T06489 (1-568)

QY 13 AAGAATGTTATCAAGGAG-----TTCATGAGGTTTAAAG 45

Db 39 LysLysLeuValLysGluGlyGluGlyTrpAspThrAlaGluThrAlaLeuLysValGlu 58

QY 46 GTTCGCATGGAAGGAACG---GTCAATGGGCACGAGTTTGAATAAGAGCGAAGGAGAG 102

Db 59 ValHisTyrThrGlyThrLeuLeuAspGlyThrLysPheAsp---SerSerArgAspArg 77

QY 103 GGGAGGCCATACGAAGGCCACATACCGTAAAGCTTAAGGTAACCAAGGGGACCTTTG 162

Db 78 GlyThrProPhe-----LysPheLysLeuGluGlnGlyGlnValile 91

QY 163 CCATTTGCTTGGGAT-----ATT 180

Db 92 Lys---GlyTrpAspGlnGlyIleLysThrMetLysLysGlyGluAsnAlaSerLeuThr 110

QY 181 TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240

Db 111 IleProProAspLeuAlaTyrGlyGluArgAla-----ProArgThrIlePro 126

QY 241 GACTATAAAAAGCTGTCTTTTCCT----- 264

Db 127 ProAsnAlaThrLeuArgPheAspValGluLeuLeuSerTrpAlaSerValLysAspIle 146

QY 265 -----GAAGGATTTAAATGGGAAAGGGTC 288

Db 147 CysLysAspGlyGlyIlePheLysValLeuValGluGlyGlnLysTrpGlu----- 164

QY 289 ATGAACCTTGAAGACGGTGGCGTGTACTGTAAACCCAGGATTCAGTTTGCAGGATGCG 348



Db 165 ---AsnProLysAspLeuAspGluValThrValLysTyrGluAlaArgLeuGluAspGly 183

QY 349 TGTTCATCTACAAGTCAAGTTCATTGGCGTGGAACCTTCCTTCCGATGGACCTGTTATG 408

Db 184 SerValSerLysSerGluSerIle----- 192

QY 409 CAAAAGAAGACAATGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTG 468

Db 193 -----GluPheSerVal-----LysAspGlyTyr 200

QY 469 TTGAAAGGAGAGATTCAAGGCTCTGAAG---CTGAAAGACGGTGGTGCATTACCTAGTT 525

Db 201 PheCysProAlaLeuSerLysAlaValLysThrMetLysLysGlyGluLysValLeuLeu 220

QY 526 GAATTCAAAAGTATTAC-----ATGGCAAGAAAGCCTGTGCGACTACCAGGG 573

Db 221 ThrValLysProGlnTyrGlyPheGlyGluGlnGlyArgAlaAlaThrGluValGluGly 240

QY 574 TACTACTATGTTGACTCCAAACTGGATATA----- 603

Db 241 AlaValProProAsnSerThrLeuHisIleAspLeuGlnLeuValSerTrpLysThrLeu 260

QY 604 ACAAGCCACAACGAGACTATACAATCGTT-----GAGCAGTAT 642

Db 261 ThrLeuIleGlyAspAspLysArgIleLeuLysLysValLeuLysGluGlyGluGlyTyr 280

QY 643 GAAAGAACCAGGAGGCCACCATC 667

Db 281 GluArg-ProAsnAspGlyAlaVal 288

RESULT 6

A36734

bacilloleptidase F (EC 3.4.21.-) precursor bpr [validated] - Bacillus subtilis

C;Species: Bacillus subtilis

C;Date: 19-May-2000 #sequence revision 19-May-2000 #text change 09-Jul-2004

C;Accession: A36734; A35131; A35750; B35750; S08223; JN0335; I39849; B69596; JU0084

R;Sloma, A.; Rufo Jr., G.A.; Rudolph, C.F.; Sullivan, B.J.; Theriault, K.A.; Pero, J.

J. Bacteriol. 172, 5520-5521, 1990

A;Reference number: A36734; MUID:90368623; PMID:2118514

A;Contents: erratum

A;Accession: A36734

A;Molecule type: DNA

A;Residues: 1-1433 <SLO>

A;Cross-references: UNIPROT:P16397; GB:M29035; NID:g143307; PIDN:AAA62679.1; PID:g143308

R;Sloma, A.; Rufo Jr., G.A.; Rudolph, C.F.; Sullivan, B.J.; Theriault, K.A.; Pero, J.

J. Bacteriol. 172, 1470-1477, 1990

A;Title: Bacilloleptidase F of Bacillus subtilis: purification of the protein and cloning

A;Reference number: A35131; MUID:90170864; PMID:2106512

A;Accession: A35131

A;Molecule type: DNA

A;Residues: 1-365,'S',367-682,'EIMP',893,'Q',895-896 <SL2>

A;Cross-references: GB:M29035

A;Note: the authors translated the codon GAA for residue 545 as Leu

R;Wu, X.C.; Nathoo, S.; Pang, A.S.H.; Carne, T.; Wong, S.L.

J. Biol. Chem. 265, 6845-6850, 1990

A;Title: Cloning, genetic organization, and characterization of a structural gene encoding

A;Reference number: A35750; MUID:90216713; PMID:2108961

A;Accession: A35750

A;Molecule type: DNA

A;Residues: 1-392,'V',394-828,'NIRTRLYSCLKFCSRHKSV' <WUA>

A;Cross-references: GB:J05400; NID:g142607; PIDN:AAA83362.1; PID:g142609

A;Note: this sequence has been corrected

A;Accession: B35750

A;Molecule type: DNA

A;Residues: 876-935,'CG' <WU2>

A;Cross-references: GB:J05400; NID:g142607; PIDN:AAA83363.1; PID:g1119197

A;Note: this sequence has been corrected

R;Masuda, E.S.; Anaguchi, H.; Sato, T.; Takeuchi, M.; Kobayashi, Y.

Nucleic Acids Res. 18, 657, 1990

A;Title: Nucleotide sequence of the sporulation gene spoIIIGA from Bacillus subtilis.

A;Reference number: S08223; MUID:90174995; PMID:2106671

A;Accession: S08223

A;Molecule type: DNA

A;Residues: 1410-1433 <MAS>

A;Cross-references: EMBL:X17344; NID:g40165; PIDN:CAA35224.1; PID:g809661

R;Kato, T.; Yamagata, Y.; Arai, T.; Ichishima, E.

Biosci. Biotechnol. Biochem. 56, 1166-1168, 1992

A;Title: Purification of a new extracellular 90-kDa serine proteinase with isoelectric p

A;Reference number: JN0335; MUID:93005071; PMID:1368833

A;Accession: JN0335

A;Molecule type: protein

A;Residues: 195-218,'A' <KAT>

A;Note: source of this material was Bacillus subtilis (natto)

R;Beall, B.; Lowe, M.; Lutkenhaus, J.

J. Bacteriol. 170, 4855-4864, 1988

A;Title: Cloning and characterization of Bacillus subtilis homologs of Escherichia coli

A;Reference number: I39846; MUID:89008108; PMID:3139638

A;Accession: I39849

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-211 <RES>

A;Cross-references: GB:M22630; NID:g142938; PIDN:AAA22458.1; PID:g551705

R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter

C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho

A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A;Authors: Foulger, D.; Fritzt, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadie, Y.; Sato, T.; Scanlon,

A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K

A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A;Reference number: A69580; MUID:98044033; PMID:9384377

A;Accession: B69596

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-1433 <KUN>

A;Cross-references: GB:Z99111; GB:Z99112; GB:AL009126; NID:g2633902; PIDN:CAB13404.1; PI

A;Experimental source: strain 168

C;Genetics:

A;Gene: bpr; bpf

A;Map position: 135 (degrees)

C;Superfamily: bacilloleptidase F; subtilisin homology

C;Keywords: extracellular protein; hydrolase; serine proteinase

F;1-30/Domain: signal sequence #status predicted <SIG>

F;31-194/Domain: propeptide #status predicted <PRO>

F;195-1433/Product: bacilloleptidase F #status experimental <MAT>

F;218-466/Domain: subtilisin homology <SBT>

F;227,274,452/Active site: Asp, His, Ser #status predicted

Alignment Scores:

Pred. No.:	3.97	Length:	1433
Score:	89.00	Matches:	56
Percent Similarity:	35.19%	Conservative:	26
Best Local Similarity:	24.03%	Mismatches:	71
Query Match:	7.33%	Indels:	80
DB:	1	Gaps:	14

US-10-006-922A-11 (1-678) x A36734 (1-1433)

QY 49 CGCATGGAAGGAACGGTCAATGGGCACGAGTTTGAAATAGAGCGGAAGGAGGAGGAGG 108

Db 521 LysAlaGluGlyGlnValSer-----ValGluGlyAspAspGlnGluPro 535

QY 109 CCATACGAAGGCCACAAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATT 168

Db 536 ProValTyrGlnHisGluLysVal---ThrGluAlaTyrGluGlyGlySerLeuProLeu 554

QY 169 GCT----- 171

Db 555 ThrLeuThrAlaGluAspAsnValSerValThrSerValLysLeuSerTyrLysLeuAsp 574







Score:	85.50	Matches:	46
Percent Similarity:	33.97%	Conservative:	25
Best Local Similarity:	22.01%	Mismatches:	59
Query Match:	7.04%	Indels:	79
DB:	2	Gaps:	10

US-10-006-922A-11 (1-678) x G69397 (1-456)

QY	76	GAGTTTGAATAGAGCGGAGGAGGCATACGAGGCCACATACCGTAAAG	135
Db	67	GluPheGluGluLeuGlyGlu--GluGlyArgPheTyrGluSerArg-----	81
QY	136	CTTAAGGTAACCAAGGGGGACCTTTGCCATTTCGTTGGGATATTTTGTCAACCACAATTT	195
Db	82	---TyrIleThrLysAspGlyArgValArgHisValTrpGlyPheThrAlaLysPheThr	100
QY	196	CAGTATGGAACCAAGGTATATGTCAAGCACCTCGCGACATACCAGACTATAAAAAGCTG	255
Db	101	PheLeuGlyLysSerTyrIleIleGlyAsnTrpIleAspValThrLysSerLysGluLeu	120
QY	256	TCATTTCCTGAAGGATTTAAATGGGAAAGGTCATGAACTTTGAAGACGGTGGCGTCGTT	315
Db	121	-----GluGlnAlaLeuArgGluSerGluGluPheTyrArg	132
QY	316	ACTGTAACCCAGGATTCC-----AGTTTGAGGATGGCTGTTTCATCTAC	360
Db	133	ThrLeuValGluAspSerLeuThrProValTyrLeuLeuGlnAspGlyIleMetValTyr	152
QY	361	-----AAGGTCAAGTTCATTGGCGTGAAC	384
Db	153	ValAsnLysAlaPheGluGluAlaThrGlyTyrLysArgGluGluIleValGlyArgAsn	172
QY	385	-----TTTCCTCCGATGGACCTGTTATGCCAAAAGAAG-----	417
Db	173	ProphePheLeuIleHisProGluAspArgGlyLeuValTyrLysArgTyrIleGluArg	192
QY	418	-----ACAATGGCGTGGGAGCCAGCACTGAGCGTTTG	450
Db	193	GluLysGlyLeuArgAspThrMetGluThrTyrSerTrp-----Argile	207
QY	451	TATCCTCGTGATGGCGTGTGAAA-----	474
Db	208	IleArgLysAspGlyGluValArgTrpValThrAlaArgProGlyArgValThrTyrArg	227
QY	475	-----GGAGAGATTTCATAAGGCT---CTG	495
Db	228	GlyArgProAlaValAlaAlaThrValValAspThrThrGluIleHisLysLeuAsnIle	247
QY	496	AAGCTGAAAGACGGTGCATTACCTA	522
Db	248	GluLeuLysLysLysGlyGluTyrLeu	256

RESULT 11  
E86445  
hypothetical protein F3C3.9 [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C;Accession: E86445  
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.P.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A;Reference number: A86141; MUID:21016719; PMID:11130712  
A;Accession: E86445  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1206 <STO>

A;Cross-references: UNIPROT:Q9FVQ7; GB:AE005172; NID:g10801373; PIDN:AAG23445.1; GSPDB:G			
C;Genetics:			
A;Map position: 1			
Alignment Scores:			
Pred. No.:	8.27	Length:	1206
Score:	85.50	Matches:	37
Percent Similarity:	41.61%	Conservative:	20
Best Local Similarity:	27.01%	Mismatches:	41
Query Match:	7.16%	Indels:	39
DB:	2	Gaps:	6
US-10-006-922A-11 (1-678) x E86445 (1-1206)			
QY 672	GAACAGATGGTGGCGTCCCTCGGTTCTTTTCATACTGCTCAACGATTGTATAGTCTTCGTT	613	
Db 997	GluleuTyrValValThrIleGlySerPheIleAspLeu-----CysIleGluPheLeu	1014	
QY 612	GTGGCTTGTATATCCAGTTTGGAGTCAAC---ATAGTAGTACCCTGGTAGCTGCACAGG	556	
Db 1015	TyrSerThrHisLeuLysPheValAsnGlyValLeuAsnPro-----	1029	
QY 555	CTTCTTTGCCATGTAAATACTTTTGAATTCAACTAGGTAATGACCACCGTCTTTTCAGCTT	496	
Db 1030	-----SerHisMetAsnAspPheGluHisSerGlyMetLeuLeuMetPhePheIleLeu	1047	
QY 495	CAGAGCCTTATGAATCTCTCCTTTCAACACGCCCATCACGAGGATACAAACGCTC-----	442	
Db 1048	GlyPheIleAlaLeuLeuSerGluLys-----ThrArgLeuLeuProLeuProGln	1064	
QY 441	-----AGTGCTGGCTTCCCAGCCCATGTCTTCTTTTGCATAAC	403	
Db 1065	GluAlaLeuCysLeuIleAlaAlaThrAlaPheThrAlaGluCysLeuLeuPhePhe	1084	
QY 402	AGGTCCATCGAAGGAAAGTTTCACGCCAATGAACCTTGACCTTGTAGATGAAACAGCCATC	343	
Db 1085	HisSerThrSerHisLys-----	1090	
QY 342	CTGCAAACTGGAATCCTGGGTTACAGTAACGACGCCACCGTCTTCAAAGTT	292	
Db 1091	-----GlyLeuGluGlyTyrTyrHisLeuLeuValPheLeuIle	1104	
RESULT 12			
F69442			
hypothetical protein AF1543 - Archaeoglobus fulgidus			
C;Species: Archaeoglobus fulgidus			
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004			
C;Accession: F69442			
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson,			
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F."			
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.			
Nature 390, 364-370, 1997			
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.N			
Smith, H.O.; Woese, C.R.; Venter, J.C.			
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo			
A;Reference number: A69250; MUID:98049343; PMID:9389475			
A;Accession: F69442			
A;Status: preliminary; nucleic acid sequence not shown; translation not shown			
A;Molecule type: DNA			
A;Residues: 1-271 <KLE>			
A;Cross-references: UNIPROT:O28729; GB:AE000996; GB:AE000782; NID:g2689319; PIDN:AAB8971			
C;Superfamily: Pyrococcus horikoshii hypothetical protein PH1420			
Alignment Scores:			
Pred. No.:	6.64	Length:	271
Score:	85.00	Matches:	44
Percent Similarity:	36.05%	Conservative:	18
Best Local Similarity:	25.58%	Mismatches:	64
Query Match:	7.00%	Indels:	46
DB:	2	Gaps:	9
US-10-006-922A-11 (1-678) x F69442 (1-271)			



QY 190 CAATTTTCAGTATGGAAGCAAGGTATATATGTCAAGCACCTT-----GCCGAC 234  
Db 65 GlnThrGluIleAsnSerLeuPheTyrThrLeuHisSerLeuAsnLeuLeuGlyGluAsp 84  
QY 235 ATACCAGACTATAAAAG-----CTGTCAATTCCTGAAGGATTTAAATGGGAAAGGGTC 288  
Db 85 LeuProAspTyrSerSerPheLeuLeuLysArgLeuGluGlyLeuLysAlaGluArgLys 104  
QY 289 ATGAACCTTTGAAGACGGTGGCGTCGTT-----ACTGTAACCCAGGATTCAGTTTG 339  
Db 105 TyrLeuLeuSerAspGlyValThrAlaThrTyrThrPheLeuGlnProAsnAlaLeu 124  
QY 340 CAGGATGGCTGTTTCATCTACAAGGTCAAGTTCATTGGCGTGAACCTTTCCTCCGATGGA 399  
Db 125 ArgAspAla-----TyrMetIleSerThrLeuLeuHisLeuTyrAsnArgAspVal 141  
QY 400 CCTGTTATGCAAAAGACAAATGGGCTGGGAAGCCAGCAGCTAGCGGTTTGTATCCTCGT 459  
Db 142 ProGluGluThrLysMetLeuVal-----ArgArgTyrArgArg 154  
QY 460 GAT-----GGCGTGTGAAAGGAGAGATTTCATAAGGCTCTGAAG 498  
Db 155 GluThrGluPheGlyValGlyTyrGlyValLysLysProAsnLeuGluGluThrPheTyr 174  
QY 499 CTGAAAGACGGTGGTCAATACCTAGTTGAATTCAAAAGTATTAC-----543  
Db 175 -----AlaSerTyrIleLeuArgAspLysAlaValIleSerPheValLysSer 190  
QY 544 -----ATGGCAAAGAAGCCCTGTGCAGCTACCAGGGTAC-----576  
Db 191 PheGluSerAsnGlyGlyPheAlaLysGlnProGlyGlyTyrProProTyrLeuGluAsp 210  
QY 577 ---TACTATGTTGACTCCAAACTGGATATAACAAGC 609  
Db 211 ThrTyrTyrAlaThrSerThrLeuSerLeuLeuSer 222  
RESULT 13  
A40351  
adhesion-type protein ADMLX - human  
C;Species: Homo sapiens (man)  
C;Date: 24-Jan-1992 #sequence\_revision 24-Jan-1992 #text\_change 15-Mar-2004  
C;Accession: A40351  
R;Legouis, R.; Hardelin, J.P.; Levilliers, J.; Claverie, J.M.; Compain, S.; Wunderle, V.  
, G.; Weissenbach, J.; Petit, C.  
Cell 67, 423-435, 1991  
A;Title: The candidate gene for the X-linked Kallmann syndrome encodes a protein related  
A;Reference number: A40351; MUID:92005720; PMID:1913827  
A;Accession: A40351  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-679 <LEG>  
A;Cross-references: GB:S60085; NID:g237595; PIDN:AAB20108.1; PID:g237597  
C;Genetics:  
A;Gene: GDB:KAL1; KAL  
A;Cross-references: GDB:120116; OMIM:308700  
A;Map position: Xp22.32-Xp22.32  
F;129-175/Domain: antileukoproteinase repeat homology <ALP>  
Alignment Scores:  
Pred. No.: 9.09 Length: 679  
Score: 84.50 Matches: 48  
Percent Similarity: 38.20% Conservative: 20  
Best Local Similarity: 26.97% Mismatches: 75  
Query Match: 6.95% Indels: 35  
DB: 2 Gaps: 10  
US-10-006-922A-11 (1-678) x A40351 (1-679)  
QY 136 CTTAAGGTAACCAAGGGGGA-----CCTTGCCATT-----168  
Db 406 ValLysThrArgLysGlyGlyIleGlnThrGlnLeuProPheGlnArgArgArgProThr 425  
QY 169 ---GCTTGGGATATTGTCACCACCAATTTTCAGTATGAAGC-----AAGGTATAT 216

Db 426 ArgProLeuGluValGlyAlaProPheTyrGlnAspGlyGlnLeuGlnValLysValTyr 445  
QY 217 GTCAAGCACCTCGCCGAC-----ATACCAGACTATAAAAAGCTGTCTCCTGAAGGA 270  
Db 446 TrpLysLysThrGluAspProThrValAsnArgTyrHisValArgTrpPheProGluAla 465  
QY 271 TTTAAATGGGAAAGGTCATGAACCTTTGAAGACGGTGGCGTCGTTACTGTAAACCCAGGAT 330  
Db 466 CysAlaHisAsnArgThrThrGlySerGluAlaSerSerGlyMetThrHisGluAsnTyr 485  
QY 331 TCCAGTTTGCAGGATGGCTGTTTCATCTACAAGGTCAAAGTTCATTGGCGTGAACCTTTCCT 390  
Db 486 IleIleLeuGlnAspLeuSerPheSerCysLysTyrLysVal-----499  
QY 391 TCCGATGGACCTGTATTGCAAAAGAAAGACAATGGGCTGGGAAGCCAGCAGCTAGCGTTTG 450  
Db 500 ThrValGlnProIleArgProLysSer-----HisSerLysAlaGluAlaVal 515  
QY 451 TAT-----CCTCGTGATGGCGTGTGAAAGGAGAGATTTCATAAGGCTCTGAAG---498  
Db 516 PhePheThrThrProProCysSerAlaLeuLysGlyLysSerHisLysProIleGlyCys 535  
QY 499 CTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATTTCATATGGCAAAGAGCCT 558  
Db 536 LeuGlyGluAlaGlyHisValLeuSerLys-----ValLeuAlaLysPro 550  
QY 559 GTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGATATAACAAGGCAC 612  
Db 551 GluAsnLeuSerAlaSerPheIleValGlnAsp---ValAsnIleThrGlyHis 567  
RESULT 14  
IJMSNG  
neural cell adhesion molecule 1 precursor, GPI-anchored splice form - mouse  
N;Alternate names: NCAM-120  
C;Species: Mus musculus (house mouse)  
C;Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 09-Jul-2004  
C;Accession: A29673; S00382; A44290  
R;Barthels, D.; Santoni, M.J.; Wille, W.; Ruppert, C.; Chaix, J.C.; Hirsch, M.R.; Fontes  
EMBO J. 6, 907-914, 1987  
A;Title: Isolation and nucleotide sequence of mouse NCAM cDNA that codes for a Mr 79,000  
A;Reference number: A29673; MUID:87246524; PMID:3595563  
A;Molecule type: mRNA  
A;Residues: 1-725 <BAR>  
A;Cross-references: UNIPROT:P13594; EMBL:Y00051; NID:g53342; PIDN:CAA68263.1; PID:g53343  
R;Barbas, J.A.; Chaix, J.C.; Steinmetz, M.; Goridis, C.  
EMBO J. 7, 625-632, 1988  
A;Title: Differential splicing and alternative polyadenylation generates distinct NCAM t  
A;Reference number: S00382; MUID:88283628; PMID:3396534  
A;Accession: S00382  
A;Molecule type: DNA  
A;Residues: 642-656, 'D', 658-725 <BA2>  
A;Cross-references: EMBL:X07195  
R;Rougon, G.; Marshak, D.R.  
J. Biol. Chem. 261, 3396-3401, 1986  
A;Title: Structural and immunological characterization of the amino-terminal domain of m  
A;Reference number: A44290; MUID:86140120; PMID:3512556  
A;Accession: A44290  
A;Molecule type: protein  
A;Residues: 20-36 <ROU>  
C;Comment: NCAM mediates cell-cell adhesion via homophilic binding with another NCAM mol  
C;Comment: Several forms of NCAM are produced by alternative splicing. See also PIR:IJMS:  
C;Genetics:  
A;Gene: NCAM  
A;Map position: 9  
A;Introns: 701/1  
C;Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; immu  
C;Keywords: alternative splicing; cell adhesion; duplication; heparin binding; membrane  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;34-98/Domain: immunoglobulin homology <IMM1>  
F;132-191/Domain: immunoglobulin homology <IMM2>  
F;152-156/Region: heparin binding #status predicted



F;161-165/Region: heparin binding #status predicted  
F;228-290/Domain: immunoglobulin homology <IMM3>  
F;263-272/Region: NCAM binding #status predicted  
F;323-388/Domain: immunoglobulin homology <IMM4>  
F;420-482/Domain: immunoglobulin homology <IMM5>  
F;519-596/Domain: fibronectin type III repeat homology <FN3A>  
F;625-685/Domain: fibronectin type III repeat homology <FN3B>  
F;41-96,139-189,235-288,330-386,427-480/Disulfide bonds: #status predicted  
F;222,316,348,424,450,479/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:  
Pred. No.: 9.22 Length: 725  
Score: 84.50 Matches: 37  
Percent Similarity: 43.79% Conservative: 30  
Best Local Similarity: 24.18% Mismatches: 57  
Query Match: 6.95% Indels: 29  
DB: 1 Gaps: 9

US-10-006-922A-11 (1-678) x IJMSNG (1-725)

QY 271 TTAAATGGGAAGGTCATGAACCTTTGAAGACGGTGGCGTCGTTACTGTA----- 321  
Db 548 PheThrTrpTyrAspAlaLysGluAlaAsnMetGluGlyIleValThrIleMetGlyLeu 567  
QY 322 ACCAGGATTCCAGTTGTCAGGATGGCTGTTTCATCTACAAGGTCAAGTTCATTGGCGTG 381  
Db 568 LysProGluThrThrTyrSerAsp-----ArgLeuAlaAlaLeu 580  
QY 382 AACTTTCCTCCGATGGACCTGTTATGCAAAAGACAATGGGCTGGGAAGCCAGCACT 441  
Db 581 AsnGlyLysGlyLeuGlyGluIleMetGlnProSerGluSerLysThrGlnProValPro 600  
QY 442 GAGCGTTTGATCCTCGTGATGGCGTGTGAAGGAGAGATTCAATAG----- 489  
Db 601 GluLeuSerAlaProLys-----LeuGluGlyGlnMetGlyGluAspGlyAsnSer 617  
QY 490 -----GCTCTGAAGCTGAAGACGGTGGT-----CATTACCTAGTT 525  
Db 618 IleLysValAsnLeuIleLysGlnAspAspGlyGlySerProIleArgHisTyrLeuVal 637  
QY 526 GAATTCAAAAGTATTACATGGCAAGACGCT--GTGCAGCTACCA---GGGTACTAC 579  
Db 638 LysTyrArgAlaLeuAlaSerGluTrpLysProGluIleArgLeuProSerGlySerHis 657  
QY 580 TATGTTGACTCCAAA---CTGGATATAACAAGCCACAAGACTATACAATCGTTGAG 636  
Db 658 HisValMetLeuLysSerLeuAspTrpAsnAlaGluTyrGluValTyrValAlaGlu 677  
QY 637 ---CAGTATGAAGAACCGAGGACGCCACCATCTGTTTC 672  
Db 678 AsnGlnGlnGlyLysSerLysAlaAlaHisPheValPhe 690

RESULT 15  
AC0314  
probable membrane protein YPO2573 [imported] - Yersinia pestis (strain CO92)  
C;Species: Yersinia pestis  
C;Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C;Accession: AC0314  
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001  
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A;Reference number: AB0001; MUID:21470413; PMID:11586360  
A;Accession: AC0314  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-2004 <KUR>  
A;Cross-references: UNIPROT:Q8ZDJ2; GB:AL590842; PIDN:CAC91375.1; PID:g15980562; GSPDB:G C;Genetics:  
A;Gene: YPO2573

Alignment Scores:

Pred. No.: 11.5 Length: 2004  
Score: 84.50 Matches: 57  
Percent Similarity: 31.52% Conservative: 30  
Best Local Similarity: 20.65% Mismatches: 81  
Query Match: 6.95% Indels: 108  
DB: 2 Gaps: 9

US-10-006-922A-11 (1-678) x AC0314 (1-2004)

QY 55 GAAGGAACGGTCAATGGGCACGAGTTTGAATAAGAGCGAGAGAGAGGGAGGCCATAC 114  
Db 905 GluAsnArgHisAsnSerAspGlyPheGluThrGluLeuGluGlnThrThrAspLeu 924  
QY 115 GAAGGCCACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTGCTTGG 174  
Db 925 GlnGlyMetAlaThrIleProLeuAspLeuLysSerTyrAlaAspAlaThrTyrGlnLeu 944  
QY 175 GATATTTGTCCACCACAATTTTCAG----- 198  
Db 945 GlnLeuLeuSerGluAlaPheGluAlaGlyGlyGlyArgSerValAlaAlaThrAlaArg 964  
QY 199 -----TATGGAAGCAAGGTATATGTCAAGCACCTTCCGACATA----- 237  
Db 965 ValLeuValSerProTyrAspSerLeuValGlyValLysAlaAspGlyAspLeuSerTyr 984  
QY 238 -----CCAGACTATAAAAAG 252  
Db 985 IleAsnArgAspAlaValArgLysLeuAsnIleAlaValAspProSerLeuAsnLys 1004  
QY 253 CTGTCATTTCCTGAA-----GGATTTAATGGGAAAGGGTCATGAAC 294  
Db 1005 IleAlaLeuProAspLeuSerLeuSerLeuIleGluGlnLysTyrIleSerValLeuThr 1024  
QY 295 TTTGAAGACGGTGGCGTC----- 312  
Db 1025 LysGlnAspSerGlyValTyrLysTyrGlnSerArgLeuLysGluGlnLeuValSerGlu 1044  
QY 313 -----GTTACTGTAAACCCAGGATTCAGT 336  
Db 1045 GlnProLeuGlnIleSerProThrGlyThrAspPheThrLeuValThrGln----- 1061  
QY 337 TTGCAGGATGGCTGTTTCATCTACAAGGTCAAGTTCATTGGCGTGAACCTTTCCCTCCGAT 396  
Db 1062 ---GlnProGlyAspPheIleLeuValLys-----AspSerGln 1074  
QY 397 GGACCTGTTATGCAAAAG-----AAGACAATGGGCTGGGAAGCCAGCACTGAGCGTTTG 450  
Db 1075 GlyGlnValLeuAsnArgIleSerTyrThrValAlaGlyAsnAlaAsnLeuThrArgSer 1094  
QY 451 TATCCTCGTGATGGCGTGTGTTGAAAGGAGAGATTTCATAAGGCTCTGAAGCTGAAA----- 504  
Db 1095 LeuAspArgAsnThrGluLeuLysLeuLysLeuAsnGlnAlaGluTyrLeuGlnGlyGlu 1114  
QY 505 -----GACGGTGGTCAATTACCTAGTTGAA 528  
Db 1115 GluIleGluIleAlaIleAsnAlaProTyrAlaGlySerGly-LeuIleThrIleGluLy 1134  
QY 529 TTCAAAAGTATTTCATGCGCAAAGAACCTGTGCAGCTACCAGGGTACTACTATGTTGAC 588  
Db 1134 sAspLysValTyrSerTrpGln----- 1141  
QY 589 TCCAAACTGGATATAACAAGCCACAAGCAAGACTATACAATCGTTG 634  
Db 1142 -----TrpPheHisSerAspThrThrSerSerValGlnArgile 1154

Search completed: June 30, 2005, 16:05:56  
Job time : 66.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: June 30, 2005, 15:49:45 ; Search time 161.5 Seconds  
(without alignments)  
4299.563 Million cell updates/sec

Title: US-10-006-922A-11  
Perfect score: 1215  
Sequence: 1 atgaggtcttccaagaatgt.....gccaccatctgttcctttaa 678

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US10006922/runat\_30062005\_110551\_9411/app\_query.fasta\_1.839  
-DB=Uniprot -QFMT=fastan -SUFFIX=rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10006922 @CGN\_1\_1\_244 @runat\_30062005\_110551\_9411 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Uniprot\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1214	99.9	225	Q9U6Y8	Q9u6y8 discosoma s
2	1196	98.4	236	Q66ND7	Q66nd7 discosoma s
3	1193	98.2	225	Q6KPF8	Q6kpf8 discosoma s
4	1191	98.0	236	Q66ND8	Q66nd8 discosoma s
5	1085.5	89.3	230	Q9GTJ7	Q9gtj7 discosoma s
6	804	66.2	221	Q66PV1	Q66pv1 acropora te
7	798	65.7	221	Q66PU8	Q66pu8 acropora ac
8	797	65.6	221	Q66PV0	Q66pv0 acropora mi
9	793	65.3	221	Q95P04	Q95p04 gonopora t
10	793	65.3	221	Q66PU9	Q66pu9 acropora hy
11	751	61.8	219	P83690	P83690 montipora e
12	745	61.3	227	Q66ND6	Q66nd6 discosoma s
13	729.5	60.0	232	Q9U6Y7	Q9u6y7 discosoma s
14	690	56.8	225	Q963F5	Q963f5 montastraea
15	684	56.3	225	Q66ND3	Q66nd3 montastraea
16	681	56.0	225	Q95UA7	Q95ua7 montastraea

17	681	56.0	225	2	Q7Z0W4	Q7z0w4 montastraea
18	679.5	55.9	227	2	Q7Z0W6	Q7z0w6 montastraea
19	678.5	55.8	227	2	Q962P9	Q962p9 montastraea
20	678.5	55.8	227	2	Q7Z0W8	Q7z0w8 montastraea
21	678	55.8	225	2	Q7Z0W5	Q7z0w5 montastraea
22	672.5	55.3	227	2	Q66ND2	Q66nd2 montastraea
23	670.5	55.2	227	2	Q66ND5	Q66nd5 montastraea
24	669.5	55.1	234	2	Q7Z0W7	Q7z0w7 montastraea
25	668.5	55.0	234	2	Q8T5F2	Q8t5f2 montastraea
26	667	54.9	266	2	Q9U6Y3	Q9u6y3 clavularia
27	663	54.6	226	2	Q8T6U0	Q8t6u0 dendroneph
28	660.5	54.4	227	2	Q66ND4	Q66nd4 montastraea
29	655.5	54.0	226	2	Q66PU5	Q66pu5 agaricia fr
30	655.5	54.0	227	2	Q95VT0	Q95vt0 montastraea
31	654.5	53.9	234	2	Q8MU47	Q8mu47 montastraea
32	652.5	53.7	225	2	Q7Z0W9	Q7z0w9 montastraea
33	642.5	52.9	225	2	Q6USK3	Q6usk3 montastraea
34	639	52.6	224	2	Q8MU48	Q8mu48 montastraea
35	631.5	52.0	225	2	Q8T5F1	Q8t5f1 montastraea
36	620.5	51.1	230	2	Q66PW1	Q66pw1 scolymia cu
37	610	50.2	223	2	Q6R8F5	Q6r8f5 astrangia l
38	608.5	50.1	231	2	Q8ISF8	Q8isf8 parasicyoni
39	599.5	49.3	225	2	Q8I6J8	Q8i6j8 trachyphyll
40	580.5	47.8	219	2	Q6R8F4	Q6r8f4 astrangia l
41	579.5	47.7	259	2	Q8MMA2	Q8mma2 agaricia fr
42	574	47.2	228	2	Q9GPI6	Q9gpi6 anemonia su
43	573.5	47.2	235	2	Q8T5F0	Q8t5f0 scolymia cu
44	570.5	47.0	231	2	Q8T6T8	Q8t6t8 discosoma s
45	566.5	46.6	239	2	Q8MMA1	Q8mmal agaricia ag

ALIGNMENTS

RESULT 1  
Q9U6Y8 PRELIMINARY; PRT; 225 AA.  
AC Q9U6Y8;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Fluorescent protein FP583.  
OS Discosoma sp.  
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;  
OC Discosomatidae; Discosoma.  
OX NCBI\_TaxID=86600;  
RN [1]\_TaxID=86600;  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99436614; PubMed=10504696;  
RA Matz M.V., Fradkov A.F., Labas Y.A., Savitsky A.P., Zaraisky A.G.,  
RA Markelov M.L., Lukyanov S.A.;  
RT "Fluorescent proteins from nonbioluminescent Anthozoa species.";  
RL Nat. Biotechnol. 17:969-973(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Matz M.V., Fradkov A.F., Labas Y.A., Savitsky A.P., Zaraisky A.G.,  
RA Markelov M.L., Lukyanov S.A.;  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF168419; AAF03369.1; -  
DR PDB; 1G7K; X-ray; A/B/C/D=1-225.  
DR PDB; 1GGX; X-ray; A/B/C/D=1-225.  
DR GO; GO:0006091; P:energy pathways; IEA.  
DR InterPro; IPR009017; GFP\_like.  
DR InterPro; IPR011584; GFP\_related.  
DR InterPro; IPR000786; Green\_fl\_protein.  
DR Pfam; PF01353; GFP; 1.  
DR PRINTS; PR01229; GFPUCRESENT.  
DR PRODom; PD013756; Green fl protein; 1.  
SQ SEQUENCE 225 AA; 25931 MW; FBF9A5369778F689 CRC64;

Alignment Scores:  
Pred. No.: 1.59e-108 Length: 225  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.92% Indels: 0  
DB: 2 Gaps: 0

US-10-006-922A-11 (1-678) x Q9U6Y8 (1-225)

QY 1 ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTCCGATGGAAGGA 60  
Db 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
QY 61 ACGGTCAATGGGCACGAGTTTGAAATAGAACGCCGAGGAGAGGGGAGGCCATACGAAGGC 120  
Db 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTGGGATATT 180  
Db 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
QY 181 TTGTCAACCAAAATTTCAAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA 240  
Db 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
QY 241 GACTATAAAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300  
Db 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
QY 301 GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360  
Db 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
QY 361 AAGGTCAAGTTTCATTTGGCGTGAACTTTCCTCCGATGGACCTGTTATGCAAAAGAAGACA 420  
Db 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGTATGGCGTGTGAAAGGAGAG 480  
Db 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
QY 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTCCTAGTTGAATTCAAAAGTATT 540  
Db 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
QY 541 TACATGGCAAAGACCTGTGCAGCTACCGGTAATACTACTATGTTGACTCCAAACTGGAT 600  
Db 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
QY 601 ATAACAAGCCACAACGAGACTATACAATCGTTGACAGTATGAAAGAACCCGAGGGACGC 660  
Db 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220

RESULT 2

Q66ND7  
ID Q66ND7 PRELIMINARY; PRT; 236 AA.  
AC Q66ND7;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Enhanced red fluorescent protein R+.  
OS Discosoma sp. RC-2004.  
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;  
OC Discosomatidae; Discosoma.  
OX NCBI\_TaxID=289055;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Carter R.W., Gibbs P.D.L., Schmale M.C.;  
RT "Cloning of Anthozoan Fluorescent Protein Genes.";  
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY679107; AAU04444.1; -.  
DR InterPro; IPR009017; GFP\_like.

DR InterPro; IPR011584; GFP\_related.  
DR InterPro; IPR000786; Green\_fl\_protein.  
DR Pfam; PF01353; GFP; 1.  
DR PRINTS; PR01229; GFPJORESCENT.  
DR ProDom; PD013756; Green\_fl\_protein; 1.  
SQ SEQUENCE 236 AA; 27032 MW; BB3844BE06829EF0 CRC64;

Alignment Scores: 8.83e-107 Length: 236  
Pred. No.: 1196.00 Matches: 221  
Score: 99.11% Conservative: 2  
Percent Similarity: 98.22% Mismatches: 2  
Best Local Similarity: 98.44% Indels: 0  
Query Match: 2 Gaps: 0  
DB: 0

US-10-006-922A-11 (1-678) x Q66ND7 (1-236)

QY 1 ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTCCGATGGAAGGA 60  
Db 1 MetSerCysSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
QY 61 ACGGTCAATGGGCACGAGTTTGAATAGAACGCCGAGGAGGGGAGGCCATACGAAGGC 120  
Db 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
QY 121 CACAATACCGTAAAGCTTAAGGTAAACCAAGGGGGACCTTTGCCATTTGCTGGGATATT 180  
Db 41 HisAsnThrValLysLeuLysValThrLysGlyProLeuProPheAlaTrpAspIle 60  
QY 181 TTGTCAACCAAAATTTCAAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA 240  
Db 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
QY 241 GACTATAAAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300  
Db 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
QY 301 GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360  
Db 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
QY 361 AAGGTCAAGTTTCATTTGGCGTGAACTTTCCTCCGATGGACCTGTTATGCAAAAGAAGACA 420  
Db 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTATGGCGTGTGAAAGGAGAG 480  
Db 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
QY 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTCACCTAGTTGAATTCAAAAGTATT 540  
Db 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysThrIle 180  
QY 541 TACATGGCAAAGACCTGTGCAGCTACCGGTAATACTACTATGTTGACTCCAAACTGGAT 600  
Db 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
QY 601 ATAACAAGCCACAACGAGACTATACAATCGTTGACAGTATGAAAGAACCCGAGGGACGC 660  
Db 201 IleThrSerHisAsnLysAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
QY 661 CACCATCTGTTCCCTT 675  
Db 221 HisHisLeuPheLeu 225

RESULT 3

Q6KF85  
ID Q6KF85 PRELIMINARY; PRT; 225 AA.  
AC Q6KF85;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Orange fluorescent protein FP586.



OS Discosoma sp. JW-2002.  
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;  
OC Discosomatidae; Discosoma.  
OX NCBI\_TaxID=208461;  
RN {1}  
RP SEQUENCE FROM N.A.  
RA Wiedenmann J., Girod A.;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF545828; AAQ11987.1; --  
DR HSSP; P42212; 1B9C.  
DR GO; GO:0006091; P:energy pathways; IEA.  
DR InterPro; IPR009017; GFP like.  
DR InterPro; IPR011584; GFP-related.  
DR InterPro; IPR000786; Green\_fl\_protein.  
DR Pfam; PF01353; GFP; 1.  
DR PRINTS; PR01229; GFLUORESCENT.  
DR ProDom; PD013756; Green\_fl\_protein; 1.  
SQ SEQUENCE 225 AA; 25791 MW; E151D0E497AA23FA CRC64;

Alignment Scores:  
Pred. No.: 1.71e-106 Length: 225  
Score: 1193.00 Matches: 221  
Percent Similarity: 98.67% Conservative: 1  
Best Local Similarity: 98.22% Mismatches: 3  
Query Match: 98.19% Indels: 0  
DB: 2 Gaps: 0

US-10-006-922A-11 (1-678) x Q6KF85 (1-225)  
QY 1 ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGGA 60  
Db 1 MetSerCysSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
QY 61 ACGGTCAATGGGACGAGTTTGAAATAGACGCGAAGAGAGAGGGAGGCCATACGAAGGC 120  
Db 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGlyGly 40  
QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
Db 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
QY 181 TTGTCACCAACAATTCAGTATGGAAGCAAGGTATATGTCAAGCACCTCGCCGACATACCA 240  
Db 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
QY 241 GACTATAAAAAGCTGTCTTCCCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAA 300  
Db 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
QY 301 GACGGTGGCGTCGTACTGTAAACCCAGGATTCAGTTTGCCAGGATGGCTGTTTCATCTAC 360  
Db 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
QY 361 AAGGTCAAGTTCATTGGCGGTGAACCTTCCCTCCGATGGACCTGTTATGCAAAAAGAGACA 420  
Db 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
QY 421 ATGGCTGGGAAGCCAGCACTGAGCGTTTGATCCTCGTATGGCGTGTGAAAGGAGAG 480  
Db 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyAsp 160  
QY 481 ATTCTAAGGCTCTGAAGCTGAAAGACGGTGGTTCATTACCTAGTTGAATTCAAAAGTATT 540  
Db 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
QY 541 TACATGGCAAAGACCTGTGCAGCTACACGGGTACTACTATGTTGACTCCAAACTGGAT 600  
Db 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrValAspSerLysLeuAsp 200  
QY 601 ATAACAAGCCACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGGAGCGC 660  
Db 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220

QY 661 CACCATCTGTTCCTT 675  
Db 221 HisHisLeuPheLeu 225  
RESULT 4  
Q66ND8  
ID Q66ND8 PRELIMINARY; PRT; 236 AA.  
AC Q66ND8;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Red fluorescent protein R1.  
OS Discosoma sp. RC-2004.  
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;  
OC Discosomatidae; Discosoma.  
OX NCBI\_TaxID=289055;  
RN {1}  
RP SEQUENCE FROM N.A.  
RA Carter R.W., Gibbs P.D.L., Schmale M.C.;  
RT "Cloning of Anthozoan Fluorescent Protein Genes.";  
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY679106; AAU04443.1; --  
DR InterPro; IPR009017; GFP\_like.  
DR InterPro; IPR011584; GFP-related.  
DR InterPro; IPR000786; Green\_fl\_protein.  
DR Pfam; PF01353; GFP; 1.  
DR PRINTS; PR01229; GFLUORESCENT.  
DR ProDom; PD013756; Green\_fl\_protein; 1.  
SQ SEQUENCE 236 AA; 27042 MW; BB3844BCD6834EF3 CRC64;  
Alignment Scores:  
Pred. No.: 2.69e-106 Length: 236  
Score: 1191.00 Matches: 220  
Percent Similarity: 98.67% Conservative: 2  
Best Local Similarity: 97.78% Mismatches: 3  
Query Match: 98.02% Indels: 0  
DB: 2 Gaps: 0  
US-10-006-922A-11 (1-678) x Q66ND8 (1-236)  
QY 1 ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGGA 60  
Db 1 MetSerCysSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
QY 61 ACGGTCAATGGGACGAGTTTGAAATAGAAGCGAAGAGAGGGAGGCCATACGAAGGC 120  
Db 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
Db 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
QY 181 TTGTCACCAACAATTCAGTATGGAAGCAAGGTATATGTCAAGCACCTCGCCGACATACCA 240  
Db 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
QY 241 GACTATAAAAAGCTGTCTTCCCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAA 300  
Db 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
QY 301 GACGGTGGCGTCGTACTGTAAACCCAGGATTCAGTTTGCCAGGATGGCTGTTTCATCTAC 360  
Db 101 AspGlyGlyValValThrValThrGlnAspProSerLeuGlnAspGlyCysPheIleTyr 120  
QY 361 AAGGTCAAGTTCATTGGCGGTGAACCTTCCCTCCGATGGACCTGTTATGCAAAAAGAGACA 420  
Db 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
QY 421 ATGGCTGGGAAGCCAGCACTGAGCGTTTGATCCTCGTATGGCGTGTGAAAGGAGAG 480  
Db 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
QY 481 ATTCTAAGGCTCTGAAGCTGAAAGACGGTGGTTCATTACCTAGTTGAATTCAAAAGTATT 540



Db	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysThrIle	180	
QY	541	TACATGGCAAAGACCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT	600	
Db	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrValAspSerLysLeuAsp	200	
QY	601	ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGACGC	660	
Db	201	IleThrSerHisAsnLysAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220	
QY	661	CACCATCTGTTTCCTT	675	
Db	221	HisHisLeuPheLeu	225	
RESULT 5				
Q9GTJ7				
ID	Q9GTJ7	PRELIMINARY;	PRT;	230 AA.
AC	Q9GTJ7;			
DT	01-MAR-2001	(TrEMBLrel. 16, Created)		
DT	01-MAR-2001	(TrEMBLrel. 16, Last sequence update)		
DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)		
DE	Red fluorescent protein.			
GN	Name=FP593;			
OS	Discosoma sp. SSAL-2000.			
OC	Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;			
OC	Discosomatidae; Discosoma.			
OX	NCBI_TaxID=137428;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20434599; PubMed=10981720; DOI=10.1016/S0014-5793(00)01895-0;			
RA	Fradkov A.F., Chen Y., Ding L., Barsova E.V., Matz M.V.,			
RA	Lukyanov S.A.;			
RT	"Novel fluorescent protein from Discosoma coral and its mutants			
RT	possesses a unique far-red fluorescence.";			
RL	FEBS Lett. 479:127-130(2000).			
DR	EMBL; AF272711;	AAG16224.1; -.		
DR	HSSP; Q9U6Y8;	1GGX.		
DR	GO; GO:0006091;	P:energy pathways; IEA.		
DR	InterPro; IPR009017;	GFP_like.		
DR	InterPro; IPR011584;	GFP_related.		
DR	InterPro; IPR000786;	Green_fl_protein.		
DR	Pfam; PF01353;	GFP; 1.		
DR	PRINTS; PR01229;	GFLUORESCENT.		
SQ	SEQUENCE	230 AA; 26370 MW; 5215B1B436D67E51	CRC64;	
Alignment Scores:				
Pred. No.:	4.32e-96	Length:	230	
Score:	1085.50	Matches:	199	
Percent Similarity:	95.58%	Conservative:	17	
Best Local Similarity:	88.05%	Mismatches:	9	
Query Match:	89.34%	Indels:	1	
DB:	2	Gaps:	1	
US-10-006-922A-11 (1-678) x Q9GTJ7 (1-230)				
QY	1	ATGAGGTCTTCCAAGATGTTATCAAGGAGTTCATGAGGTTTAAGGTTTCGCATGGAAGGA	60	
Db	1	MetSerCysSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20	
QY	61	ACGGTCAATGGGCACGAGTTTGAAATAGAACGGCAAGGAGAGGGGAGGCCATACGAAGGC	120	
Db	21	ThrValAsnGlyHisGluPheGluIleLysGlyGluGlyGluGlyArgProTyrGluGly	40	
QY	121	CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT	180	
Db	41	HisCysSerValLysLeuMetValThrLysGlyGlyProLeuProPheAlaPheAspIle	60	
QY	181	TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA	240	
Db	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80	
QY	241	GACTATAAAAAGCTGTCAATTTCTCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAA	300	

Db	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100	
QY	301	GACGGTGGCGTCGTTACTGTAAACCCAGGATTCCAGTTTGCAGGATGGCTGTTTCATCTAC	360	
Db	101	AspGlyGlyValValThrValSerGlnAspSerSerLeuLysAspGlyCysPheIleTyr	120	
QY	361	AAGGTCAAGTTTCATTGGCGTGAACTTTCCCTCCGATGGACCTGTTATGCAAAAAGAAGACA	420	
Db	121	GluValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnArgArgThr	140	
QY	421	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG	480	
Db	141	ArgGlyTrpGluAlaSerSerGluArgLeuTyrProArgAspGlyValLeuLysGlyAsp	160	
QY	481	ATTCAATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT	540	
Db	161	IleHisMetAlaLeuArgLeuGluGlyGlyGlyHisTyrLeuValGluPheLysSerIle	180	
QY	541	TACATGGCAAAGAAGCCT--GTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTG	597	
Db	181	TyrMetValLysLysProSerValGlnLeuProGlyTyrTyrValAspSerLysLeu	200	
QY	598	GATATAACAAGCCACACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGA	657	
Db	201	AspMetThrSerHisAsnGluAspTyrThrValValGluGlnTyrGluLysThrGlnGly	220	
QY	658	CGCCACCATCTGTTCCCTT	675	
Db	221	ArgHisHisPropheile	226	
RESULT 6				
Q66PV1				
ID	Q66PV1	PRELIMINARY;	PRT;	221 AA.
AC	Q66PV1;			
DT	25-OCT-2004	(TrEMBLrel. 28, Created)		
DT	25-OCT-2004	(TrEMBLrel. 28, Last sequence update)		
DT	25-OCT-2004	(TrEMBLrel. 28, Last annotation update)		
DE	Chromoprotein.			
OS	Acropora tenuis (Purple tipped acropora).			
OC	Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;			
OC	Astrocoeniina; Acroporidae; Acropora.			
OX	NCBI_TaxID=70783;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Alieva N.O., Meleshkevitch E.A., Field S.F., Matz M.V.;			
RT	"Survey of coral GFP-like proteins.";			
RL	Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AY646074;	AAU06853.1;	-	
DR	InterPro; IPR009017;	GFP like.		
DR	InterPro; IPR011584;	GFP_related.		
DR	InterPro; IPR000786;	Green_fl_protein.		
DR	Pfam; PF01353;	GFP; 1.		
DR	PRINTS; PR01229;	GFLUORESCENT.		
DR	ProDom; PD013756;	Green_fl_protein; 1.		
SQ	SEQUENCE	221 AA;	25103 MW;	3BEA8B848C8F5DA4 CRC64;
Alignment Scores:				
Pred. No.:	7.39e-69	Length:	221	
Score:	804.00	Matches:	148	
Percent Similarity:	79.36%	Conservative:	25	
Best Local Similarity:	67.89%	Mismatches:	45	
Query Match:	66.17%	Indels:	0	
DB:	2	Gaps:	0	
US-10-006-922A-11 (1-678) x Q66PV1 (1-221)				
QY	16	AATGTTATCAAGGAGTTCATGAGGTTTAAGGTTTCGCATGGAAGGAACGGTCAATGGGCAC	75	
Db	2	SerValIleAlaLysGlnMetThrTyrLysValTyrMetSerGlyThrValAsnGlyHis	21	
QY	76	GAGTTTGAATAGAAGCGGAAGGAGAGGGAGGCCCATACGAAAGGCCACAATACCGTAAAG	135	

Db 22 TyrPheGluValGluGlyAspGlyLysGlyLysProTyrGluGlyGluGlnThrValLys 41

QY 136 CTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATTTGTCCACCACAATTT 195

Db 42 LeuThrValThrLysGlyProLeuProPheAlaTrpAspIleLeuSerProGlnPhe 61

QY 196 CAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCAGACTATAAAAAGCTG 255

Db 62 GlnTyrGlySerIleProPheThrLysTyrProGluAspIleProAspTyrValLysGln 81

QY 256 TCATTTCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAAGACGGTGGCGTCGTT 315

Db 82 SerPheProGluGlyTyrThrTrpGluArgIleMetAsnPheGluAspAlaValCys 101

QY 316 ACTGTAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTACAGGTCAAGTTCATT 375

Db 102 ThrValSerAsnAspSerIleGlnGlyAsnCysPheIleTyrHisValLysPheSer 121

QY 376 GGC GTGAACCTTCTCCGATGGACCTGTATGCAAAAGAAAGACAATGGGCTGGGAAGCC 435

Db 122 GlyValAsnPheProProAsnGlyProValMetGlnLysLysThrGlnGlyTrpGluPro 141

QY 436 AGCACTGAGCGTTTGATCCTCGTGATGGCTGTGAAAGGAGAGATTTCATAAGGCTCTG 495

Db 142 AsnThrGluArgLeuPheAlaArgAspGlyMetLeuIleGlyAsnAsnPheMetAlaLeu 161

QY 496 AAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATTACATGGCAAGAAC 555

Db 162 LysLeuGluGlyGlyGlyHisTyrLeuCysGluPheLysSerThrTyrLysAlaLysLys 181

QY 556 CCTGTGCAGCTACCAAGGGTACTACTATGTGACTCCAAACTGGATATAACAAGCCACAAC 615

Db 182 ProValArgMetProGlyTyrHisTyrValAspArgLysLeuAspValThrAsnHisAsn 201

QY 616 GAAGACTATACAATCGTTGAGCAGTATGAAGAACCGAGGGACGCCACCACCTCTG 669

Db 202 LysAspTyrThrSerValGluGlnCysGluIleSerIleAlaArgLysProLeu 219

RESULT 7

Q66PU8 PRELIMINARY; PRT; 221 AA.

ID Q66PU8

AC Q66PU8;

DT 25-OCT-2004 (TrEMBLrel. 28, Created)

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Chromoprotein.

OS Acropora aculeus.

OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;

OC Astrocoeniina; Acroporidae; Acropora.

OX NCBI\_TaxID=287157;

RN [1]

RP SEQUENCE FROM N.A.

RA Alieva N.O., Meleshkevitch E.A., Field S.F., Matz M.V.;

RT "Survey of coral GFP-like proteins.";

RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY646077; AAU06856.1; -.

DR InterPro; IPR009017; GFP\_like.

DR InterPro; IPR011584; GFP\_related.

DR InterPro; IPR000786; Green\_fl\_protein.

DR Pfam; PF01353; GFP; 1.

DR PRINTS; PR01229; GFLUORESCENT.

DR ProDom; PD013756; Green fl protein; 1.

SQ SEQUENCE 221 AA; 24976 MW; 307DE4FC4B018B43 CRC64;

Alignment Scores:

Pred. No.:	2.81e-68	Length:	221
Score:	798.00	Matches:	145
Percent Similarity:	80.48%	Conservative:	24
Best Local Similarity:	69.05%	Mismatches:	41
Query Match:	65.68%	Indels:	0
DB:	2	Gaps:	0

US-10-006-922A-11 (1-678) x Q66PU8 (1-221)

QY 16 AATGTTATCAAGGAGTTCATGAGGTTTAAGGTTCCGATGGAAGCAACGGTCAATGGGCAC 75

Db 2 SerValIleAlaLysGlnMetThrTyrLysValTyrMetSerGlyThrValAsnGlyHis 21

QY 76 GAGTTTGAAATAGAAAGCGAAGGAGAGGGGAGGCCCATACGAAGGCACAAATACCGTAAAG 135

Db 22 TyrPheGluValGluGlyAspGlyLysGlyLysProTyrGluGlyGluGlnThrValLys 41

QY 136 CTTAAGGTAACCAAGGGGGACCTTTGCCCATTGCTTGGGATATTTGTCCACCACAATTT 195

Db 42 LeuThrValThrLysGlyProLeuProPheAlaTrpAspIleLeuSerProGlnSer 61

QY 196 CAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCAGACTATAAAAAGCTG 255

Db 62 GlnTyrGlySerIleProPheThrLysTyrProAspAspIleProAspTyrValLysGln 81

QY 256 TCATTTCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAAGACGGTGGCGTCGTT 315

Db 82 SerPheProGluGlyTyrThrTrpGluArgIleMetAsnPheGluAspGlyAlaValCys 101

QY 316 ACTGTAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTACAGGTCAAGTTCATT 375

Db 102 ThrValSerAsnAspSerIleGlnGlyAsnCysPheIleTyrAsnValLysPheSer 121

QY 376 GGC GTGAACCTTCTCCGATGGACCTGTATGCAAAAGAAAGACAATGGGCTGGGAAGCC 435

Db 122 GlyLeuAsnPheProProAsnGlyProValMetGlnLysLysThrGlnGlyTrpGluPro 141

QY 436 AGCACTGAGCGTTTGATCCTCGTGATGGCTGTGAAAGGAGAGATTTCATAAGGCTCTG 495

Db 142 AsnThrGluArgLeuPheAlaArgAspGlyMetLeuIleGlyAsnAsnPheMetAlaLeu 161

QY 496 AAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATTACATGGCAAGAAC 555

Db 162 LysLeuGluGlyGlyGlyHisTyrLeuCysGluPheLysSerThrTyrLysAlaLysLys 181

QY 556 CCTGTGCAGCTACCAAGGGTACTACTATGTGACTCCAAACTGGATATAACAAGCCACAAC 615

Db 182 ProValArgMetProGlyTyrHisTyrValAspArgLysLeuAspValThrAsnHisAsn 201

QY 616 GAAGACTATACAATCGTTGAGCAGTATGAAGCAATGAA 645

Db 202 ArgAspTyrThrSerValGluGlnCysGlu 211

RESULT 8

Q66PV0 PRELIMINARY; PRT; 221 AA.

ID Q66PV0

AC Q66PV0;

DT 25-OCT-2004 (TrEMBLrel. 28, Created)

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Chromoprotein.

OS Acropora millepora (Coral).

OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;

OC Astrocoeniina; Acroporidae; Acropora.

OX NCBI\_TaxID=45264;

RN [1]

RP SEQUENCE FROM N.A.

RA Alieva N.O., Meleshkevitch E.A., Field S.F., Matz M.V.;

RT "Survey of coral GFP-like proteins.";

RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY646075; AAU06854.1; -.

DR InterPro; IPR009017; GFP\_like.

DR InterPro; IPR011584; GFP\_related.

DR InterPro; IPR000786; Green\_fl\_protein.

DR Pfam; PF01353; GFP; 1.

DR PRINTS; PR01229; GFLUORESCENT.

DR ProDom; PD013756; Green fl protein; 1.

SQ SEQUENCE 221 AA; 24987 MW; 546D2CFFAP2552F2 CRC64;

Alignment Scores:

Pred. No.:	3.51e-68	Length:	221
------------	----------	---------	-----





OC Astrocoeniina; Acroporidae; Acropora.  
OX NCBI\_TaxID=55974;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Alieva N.O., Meleshkevitch E.A., Field S.F., Matz M.V.;  
RT "Survey of coral GFP-like proteins.";  
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY646076; AAU06855.1; -  
DR InterPro; IPR009017; GFP\_like.  
DR InterPro; IPR011584; GFP\_related.  
DR Pfam; PF01353; GFP; 1.  
DR ProDom; PD013756; Green\_fl\_protein; 1.  
SQ SEQUENCE 221 AA; 24963 MW; FC9D9E749069D079 CRC64;

Alignment Scores:  
Pred. No.: 8.57e-68 Length: 221  
Score: 793.00 Matches: 144  
Percent Similarity: 80.48% Conservative: 25  
Best Local Similarity: 68.57% Mismatches: 41  
Query Match: 65.27% Indels: 0  
DB: 2 Gaps: 0

US-10-006-922A-11 (1-678) x Q66PU9 (1-221)

QY	16	AATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGCAACGGTCAATGGGCAC	75
Db	2	SerValIleAlaThrGlnMetThrTyrLysValTyrMetSerGlyThrValAsnGlyHis	21
QY	76	GAGTTTGAAATAGAGCGAAGGAGAGAGGGGAGGCATACGAAGGCCACATACCGTTAAAG	135
Db	22	TyrPheGluValGluGlyAspGlyLysGlyLysProTyrGluGlyGluGlnThrValArg	41
QY	136	CTTAAGGTAACCAAGGGGGACCTTTGGCCATTTGCTTGGGATATTTTGTCCACCACAATT	195
Db	42	LeuThrValThrLysGlyGlyProLeuProPheAlaTrpAspIleLeuSerProGlnSer	61
QY	196	CAGTATGGAAGCAAGGTATATGTCAAGCACCTTCGCGACATACCAGACTATAAAAAGCTG	255
Db	62	GlnTyrGlySerIleProPheThrLysTyrProGluAspIleProAspTyrValLysGln	81
QY	256	TCATTTCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAAGACGGTGGCTCGTT	315
Db	82	SerPheProGluGlyTyrThrTyrGluArgIleMetAsnPheGluAspGlyAlaValCys	101
QY	316	ACTGTAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTACAAGGTCAAGTTTCATT	375
Db	102	ThrValSerAsnAspSerSerIleGlnGlyAsnCysPheIleTyrHisValLysPheSer	121
QY	376	GGCGTGAACTTTCCTCCGATGGACCTGTTATGCAAAAGAAGACAATGGGCTGGGAAGCC	435
Db	122	GlyLeuAsnPheProProAsnGlyProValMetGlnLysLysThrGlnGlyTrpGluPro	141
QY	436	AGCACTGAGCGTTTGTATCCTCGTATGGCGTGTGAAAGGAGAGATTTCATAAGGCTCTG	495
Db	142	AsnThrGluArgLeuPheAlaArgAspGlyValIleGlyAsnAsnPheMetAlaLeu	161
QY	496	AAGCTGAAAGACGGTGGTCACTTACCTAGTTGAAATTCAAAAGTATTTACATGGCAAGAAG	555
Db	162	LysLeuGluGlyGlyGlyHisTyrLeuCysGluPheLysSerThrTyrLysAlaLysLys	181
QY	556	CCTGTGCAGCTACCAAGGGTACTACTATGTTGACTCCAAACTGGATATAACAAGCCACAAC	615
Db	182	ProValLysMetProGlyTyrHisPheValAspArgLysLeuAspValThrAsnHisAsn	201
QY	616	GAAGACTATACAATCGTTGAGCAGTATGAA	645
Db	202	LysAspTyrThrSerValGluGlnArgGlu	211

RESULT 11  
P83690  
ID P83690 PRELIMINARY; PRT; 219 AA.  
AC P83690;

DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Luminescent protein.  
OS Montipora efflorescens (Coral).  
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;  
OC Astrocoeniina; Acroporidae; Montipora.  
OX NCBI\_TaxID=105610;  
RN [1]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RX PubMed=12623015; DOI=10.1016/S0969-2126(03)00028-5;  
RA Prescott M., Ling M., Beddoe T., Oakley A.J., Dove S.,  
RA Hoegh-Guldberg O., Devenish R.J., Rossjohn J.;  
RT "The 2.2 a crystal structure of a pocilloporin pigment reveals a  
RT nonplanar chromophore conformation.";  
RL Structure 11:275-284(2003).  
DR PDB; 1MOU; X-ray; A=-.  
DR PDB; 1MOV; X-ray; A=1-219.  
DR GO; GO:0006091; P:energy pathways; IEA.  
DR InterPro; IPR009017; GFP\_like.  
DR InterPro; IPR011584; GFP\_related.  
DR InterPro; IPR000786; Green\_fl\_protein.  
DR Pfam; PF01353; GFP; 1.  
DR ProDom; PD013756; Green\_fl\_protein; 1.  
KW 3D-structure.  
SQ SEQUENCE 219 AA; 24569 MW; 434B2C95CB2DC7F9 CRC64;

Alignment Scores:  
Pred. No.: 9.91e-64 Length: 219  
Score: 751.00 Matches: 140  
Percent Similarity: 77.51% Conservative: 22  
Best Local Similarity: 66.99% Mismatches: 45  
Query Match: 61.81% Indels: 2  
DB: 2 Gaps: 1

US-10-006-922A-11 (1-678) x P83690 (1-219)

QY	19	GTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGCAACGGTCAATGGGCACGAG	78
Db	3	ValIleAlaThrGlnMetThrTyrLysValTyrMetSerGlyThrValAsnGlyHisTyr	22
QY	79	TTTGAATAGAGCGAAGGAGAGAGGGGAGGCATACGAAGGCCACATACCGTTAAAGCTT	138
Db	23	PheGluValGluGlyAspGlyLysGlyArgProTyrGluGlyGluGlnThrValLysLeu	42
QY	139	AAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATTTTGTCCACCACAATTTTCAG	198
Db	43	ThrValThrLysGlyGlyProLeuProPheAlaTrpAspIleLeuSerPro-----Gln	60
QY	199	TATGGAAGCAAGGTATATGTCAAGCACCTTCGCGACATACCAGACTATAAAAAGCTGTCA	258
Db	61	Cys**SerIleProPheThrLysTyrProGluAspIleProAspTyrValLysGlnSer	80
QY	259	TTTCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAAGACGGTGGCTCGTTACT	318
Db	81	PheProGluGlyPheThrTrpGluArgIleMetAsnPheGluAspGlyAlaValCysThr	100
QY	319	GTAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTACAAGGTCAAGTTTCATTGGC	378
Db	101	ValSerAsnAspSerSerIleGlnGlyAsnCysPheThrTyrHisValLysPheSerGly	120
QY	379	GTGAACCTTTCCTCCGATGGACCTGTTATGCAAAAGAAGACAATGGGCTGGGAAGCCAGC	438
Db	121	LeuAsnPheProProAsnGlyProValMetGlnLysLysThrGlnGlyTrpGluProHis	140
QY	439	ACTGAGCGTTTGTATCCTCGTATGGCGTGTGAAAGGAGAGATTTCATAAGGCTCTGAAG	498
Db	141	SerGluArgLeuPheAlaArgGlyGlyMetLeuIleGlyAsnAsnPheMetAlaLeuLys	160
QY	499	CTGAAAGACGGTGGTCACTTACCTAGTTGAATTCAAAAGTATTTACATGGCAAGAAGCCT	558
Db	161	LeuGluGlyGlyGlyHisTyrLeuCysGluPheLysThrThrTyrLysAlaLysLysPro	180



QY	559	GTGCAGCTACCAGGTTACTACTATGTTGACTCTCAAACTGGATATAACAAGCCACAACGAA	618
Db	181	ValLysMetProGlyTyrHisTyrValAspArgLysLeuAspValThrAsnHisAsnLys	200
QY	619	GACTATACAATCGTTGACGACGATGAA	645
Db	201	AspTyrThrSerValGluGlnCysGlu	209
RESULT 12			
Q66ND6			
ID	Q66ND6	PRELIMINARY;	PRT; 227 AA.
DT	25-OCT-2004	(TrEMBLrel. 28, Created)	
DT	25-OCT-2004	(TrEMBLrel. 28, Last sequence update)	
DT	25-OCT-2004	(TrEMBLrel. 28, Last annotation update)	
DE	Cyan fluorescent protein Cl.		
OS	Discosoma striata.		
OC	Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;		
OC	Discosomatidae; Discosoma.		
OX	NCBI_TaxID=105400;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Carter R.W., Gibbs P.D.L., Schmale M.C.;		
RT	"Cloning of Cnidarian Fluorescent Protein Genes.";		
RL	Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.		
DR	EMBL; AY679108; AAU04445.1; -.		
DR	InterPro; IPR009017; GFP_like.		
DR	InterPro; IPR011584; GFP_related.		
DR	InterPro; IPR000786; Green_fl_protein.		
DR	InterPro; IPR000169; Pept_cys_acsite.		
DR	Pfam; PF01353; GFP; 1.		
DR	PRINTS; PR01229; GFLUORESCENT.		
DR	ProDom; PD013756; Green_fl_protein; 1.		
DR	PROSITE; PS00639; THIOLESTERASE_HIS; UNKNOWN 1.		
SQ	SEQUENCE	227 AA; 25640 MW; 5686393306F44B244	CRC64;
Alignment Scores:			
Pred. No.:	3.8e-63	Length:	227
Score:	745.00	Matches:	135
Percent Similarity:	75.68%	Conservative:	33
Best Local Similarity:	60.81%	Mismatches:	54
Query Match:	61.32%	Indels:	0
DB:	2	Gaps:	0
US-10-006-922A-11 (1-678) x Q66ND6 (1-227)			
QY	1	ATGAGGCTCTCCAAGATGTTATCAAGGAGTTCATGAGGTTTAAGGTTTCGATGAAGGA	60
Db	1	MetSerTrpSerLysSerValIleLysGluGluMetLeuIleAsnLeuHisLeuGluGly	20
QY	61	ACGGTCAATGGGCACGAGTTTGAAATAGAAGCGCAAGGAGAGGGAGGCCATACGAAGGC	120
Db	21	ThrPheAsnGlyHisTyrPheGluIleLysGlyLysGlyAsnProAsnGluGly	40
QY	121	CACAATACCGTAAAGCTTAAGGTAACCAAGGGGACCTTTGCCATTTGCTGGGATATT	180
Db	41	ThrAsnThrValThrLeuGluValThrLysGlyGlyProLeuProPheGlyTrpHisIle	60
QY	181	TTGTCACCAACAATTTCAATGATGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA	240
Db	61	LeuCysProGlnPheGlnTyrGlyAsnLysAlaPheValHisHisProAspIlePro	80
QY	241	GACTATAAAAAGCTGTCATTTCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTGAA	300
Db	81	AspTyrLeuLysLeuSerPheProGluGlyTyrThrTrpGluArgSerMetHisPheGlu	100
QY	301	GACGGTGGCGTCTACTGTAAACCCAGGATTCAGTTTCAGGATGGCTGTTTCATCTAC	360
Db	101	AspGlyGlyLeuCysCysIleThrAsnAspIleSerLeuThrGlyAsnCysPheAsnTyr	120
QY	361	AAGGTCAAGTTTCATTTGGCGTGAACCTTTCCTCCGATGGACCTGTTATGCAAGAAGACA	420
Db	121	AspIleLysPheThrGlyLeuAsnPheProProAsnGlyProValValGlnLysLysThr	140

QY	421	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGCGGTGTTGAAAGGAGAG	480
Db	141	ThrGlyTrpGluProSerThrGluArgLeuTyrProArgAspGlyValLeuIleGlyAsp	160
QY	481	ATTCAATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT	540
Db	161	IleHisHisAlaLeuThrValAlaGlyGlyGlyHisTyrValCysAspIleLysThrVal	180
QY	541	TACATGGCAAAAGACCTGTGCAGCTACCAGGGTACTACTATGTGACTCCAAACTGGAT	600
Db	181	TyrArgAlaLys**ProValLysMetProGlyTyrHisTyrValAspProLysLeuVal	200
QY	601	ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCAGGAGCGC	660
Db	201	IleArgSerAsnAspLysGluPheMetLysValGluHisGluIleAlaValAlaArg	220
QY	661	CACCAT	666
Db	221	***His	222
RESULT 13			
Q9U6Y7			
ID	Q9U6Y7	PRELIMINARY;	PRT; 232 AA.
AC	Q9U6Y7;		
DT	01-MAY-2000	(TrEMBLrel. 13, Created)	
DT	01-MAY-2000	(TrEMBLrel. 13, Last sequence update)	
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)	
DE	Fluorescent protein FP483.		
OS	Discosoma striata.		
OC	Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;		
OC	Discosomatidae; Discosoma.		
OX	NCBI_TaxID=105400;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=99436614; PubMed=10504696;		
RA	Matz M.V., Fradkov A.F., Labas Y.A., Savitsky A.P., Zeraisky A.G.,		
RA	Markelov M.L., Lukyanov S.A.;		
RT	"Fluorescent proteins from nonbioluminescent Anthozoa species.";		
RL	Nat. Biotechnol. 17:969-973(1999).		
DR	EMBL; AF168420; AAF03370.1; -.		
DR	HSSP; Q9U6Y8; IGGX.		
DR	GO; GO:0006091; P:energy pathways; IEA.		
DR	InterPro; IPR009017; GFP_like.		
DR	InterPro; IPR011584; GFP_related.		
DR	InterPro; IPR000786; Green_fl_protein.		
DR	Pfam; PF01353; GFP; 1.		
DR	PRINTS; PR01229; GFLUORESCENT.		
DR	ProDom; PD013756; Green_fl_protein; 1.		
SQ	SEQUENCE	232 AA; 26435 MW; AA8F18EEE283CE4D	CRC64;
Alignment Scores:			
Pred. No.:	1.2e-61	Length:	232
Score:	729.50	Matches:	133
Percent Similarity:	76.00%	Conservative:	38
Best Local Similarity:	59.11%	Mismatches:	53
Query Match:	60.04%	Indels:	1
DB:	2	Gaps:	1
US-10-006-922A-11 (1-678) x Q9U6Y7 (1-232)			
QY	1	ATGAGGCTTCCCAAGAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTTCGATGAAGGA	60
Db	1	MetSerCysSerLysSerValIleLysGluGluMetLeuIleAspLeuHisLeuGluGly	20
QY	61	ACGGTCAATGGGCACGAGTTTGAAATAGAAGCGCAAGGAGAGGGAGGCCATACGAAGGC	120
Db	21	ThrPheAsnGlyHisTyrPheGluIleLysGlyLysGlyLysGlyGlnProAsnGluGly	40
QY	121	CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTGGGATATT	180
Db	41	ThrAsnThrValThrLeuGluValThrLysGlyGlyProLeuProPheGlyTrpHisIle	60

QY 181 TTGTCACCACAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240  
Db 61 LeuCysProGlnPheGlnTyrGlyAsnLysAlaPheValHisHisProAspAsnIleHis 80  
QY 241 GACTATAAAAAGCTGCTCATTTCTGAAGGATTTAAATGGGAAAGGGTCATGAACTTTGAA 300  
Db 81 AspTyrLeuLysLeuSerPheProGluGlyTyrThrTrpGluArgSerMetHisPheGlu 100  
QY 301 GACGGTGGCGTCGTACTACTGTAAACCCAGGATTCAGATTTCAGATTTCAGGATGGCTTTTCATCTAC 360  
Db 101 AspGlyGlyLeuCysCysIleThrAsnAspIleSerLeuThrGlyAsnCysPheTyrTyr 120  
QY 361 AAGGTCAAGTTTCATTGGCGTGAACTTCCTCCGATGGACCTGTTATGCAAAAGAGACA 420  
Db 121 AspIleLysPheThrGlyLeuAsnPheProAsnGlyProValvalGlnLysLysThr 140  
QY 421 ATGGCTGGGAAGCCAGCACTGAGCGTTTTCATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
Db 141 ThrGlyTrpGluProSerThrGluArgLeuTyrProArgaspGlyValleuIleGlyAsp 160  
QY 481 ATTATAAGGCTCTGAAGCTGAAAGACGGTGGTTCATTACCTAGTTGAATTCAAAAGTATT 540  
Db 161 IleHisHisAlaLeuThrValGluGlyGlyHisTyrAlaCysAspIleLysThrVal 180  
QY 541 TACATGGCAAAGAG--CCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTG 597  
Db 181 TyrArgAlaLysLysAlaAlaLeuLysMetProGlyTyrHisTyrValAspThrLysLeu 200  
QY 598 GATATAACAAGCCACAAGCAAGACTATACAATCGTTGAGCAGTATGAAGAACCAGGGA 657  
Db 201 ValIleTrpAsnAspLysGluPheMetLysValGluGluHisGluIleAlaValAla 220  
QY 658 CGCCACCATCTGTTC 672  
Db 221 ArgHisHisProphe 225  
RESULT 14  
Q963F5 PRELIMINARY; PRT; 225 AA.  
AC Q963F5;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Green fluorescent protein.  
OS Montastrea cavernosa (great star coral).  
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;  
OC Faviina; Faviidae; Montastrea.  
OX NCBI\_TaxID=63558;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lesser M.P., Barry T.M., Mazel C., Matz M.V., Lukyanov S.A.,  
RA Falkowski P., Gorbunov M., Kolber Z.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF384683; AAK62982.2; -.  
DR HSSP; Q9U6Y8; 1GGX.  
DR GO; GO:0006091; P:energy pathways; IEA.  
DR InterPro; IPR009017; GFP\_like.  
DR InterPro; IPR011584; GFP\_related.  
DR InterPro; IPR000786; Green\_fl\_protein.  
DR Pfam; PF01353; GFP; 1.  
DR PRINTS; PR01229; GFLUORESCENT.  
SQ SEQUENCE 225 AA; 25847 MW; 77DE7D7C616929AF CRC64;

Alignment Scores:  
Pred. No.: 7.95e-58 Length: 225  
Score: 690.00 Matches: 120  
Percent Similarity: 75.23% Conservative: 44  
Best Local Similarity: 55.05% Mismatches: 54  
Query Match: 56.79% Indels: 0  
DB: 2 Gaps: 0

US-10-006-922A-11 (1-678) x Q963F5 (1-225)

QY 16 AATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGATGGAAGGAACGGTCAATGGGCAC 75  
Db 2 SerValIleLysProIleMetGluIleLysLeuArgMetGlnGlyValValAsnGlyHis 21  
QY 76 GAGTTTGAATAGAAAGCGGAAGGAGAGGGAGGCCCATACGAAGGCCACATACCGTAAAG 135  
Db 22 LysPheValIleLysGlyGluGlyGluGlyLysPropheGluGlyThrGlnThrIleAsn 41  
QY 136 CTTAAGGTAACCAAGGGGGGACCTTTGCCATTTGCTGGGATATTTTGTCAACCACAATTT 195  
Db 42 LeuThrValLysGluGlyAlaProLeuProPheAlaTyrAspIleLeuThrSerAlaPhe 61  
QY 196 CAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCCAGACTATAAAAAGCTG 255  
Db 62 GlnTyrGlyAsnArgValPheThrLysTyrProAspAspIleProAspTyrPheLysGln 81  
QY 256 TCATTTCCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAAGACGGTGGCTCGTT 315  
Db 82 ThrPheProGluGlyTyrSerTrpGluArgIleMetAlaTyrGluAspGlnSerIleCys 101  
QY 316 ACTGTAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTACAGGTCGAAGTTCATT 375  
Db 102 ThrAlaThrSerAspIleLysMetGluGlyAspCysPheIleTyrGluIleGlnPheHis 121  
QY 376 GGCCTGAACCTTTCCTTCCGATGGACCTGTTATGCAAAAGAAGACAAATGGGCTGGGAAGCC 435  
Db 122 GlyValAsnPheProProAsnGlyProValMetGlnLysLysThrLeuLysTrpGluPro 141  
QY 436 AGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAGATTCATAAGGCTCTG 495  
Db 142 SerThrGluLysMetTyrValArgAspGlyValLeuLysGlyAspValAsnMetAlaLeu 161  
QY 496 AAGCTGAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATTTACATGGCAAGCAAC 555  
Db 162 LeuLeuGluGlyGlyGlyHistyrArgCysAspPheArgSerThrTyrLysAlaLysLys 181  
QY 556 CCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGATATAACAAGCCACAAC 615  
Db 182 ArgValGlnLeuProAspTyrHisPheValAspHisArgIleGluIleLeuSerHisAsp 201  
QY 616 GAAGACTATACAATCGTTGAGCAGTATGAAGAAGACCGAGGGACGCCACCATCTG 669  
Db 202 AsnAspTyrAsnThrValLysLeuSerGluAspAlaGluAlaArgTyrSerMet 219  
RESULT 15  
Q66ND3 PRELIMINARY; PRT; 225 AA.  
AC Q66ND3;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Green fluorescent protein G2.  
OS Montastrea cavernosa (great star coral).  
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;  
OC Faviina; Faviidae; Montastrea.  
OX NCBI\_TaxID=63558;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Carter R.W., Gibbs P.D.L., Schmale M.C.;  
RT "Cloning of Cnidarian Fluorescent Protein Genes.";  
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY679111; AAU04448.1; -.  
DR InterPro; IPR009017; GFP\_like.  
DR InterPro; IPR011584; GFP\_related.  
DR InterPro; IPR000786; Green\_fl\_protein.  
DR Pfam; PF01353; GFP; 1.  
DR PRINTS; PR01229; GFLUORESCENT.  
DR ProDom; PD013756; Green\_fl\_protein; 1.  
SQ SEQUENCE 225 AA; 25910 MW; A744E8D6A6C423AE CRC64;

Alignment Scores:  
Pred. No.: 3.02e-57 Length: 225  
Score: 684.00 Matches: 120

Percent Similarity: 73.85%      Conservative: 41  
Best Local Similarity: 55.05%      Mismatches: 57  
Query Match: 56.30%      Indels: 0  
DB: 2      Gaps: 0

US-10-006-922A-11 (1-678) x Q66ND3 (1-225)

QY	16	AATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGGAACGGTCAATGGGCAC	75
Db	2	SerValIleLysProAspMetLysIleLysLeuArgMetGluGlyAlaValAsnGlyHis	21
QY	76	GAGTTTGAAATAGAACGCGAAGAGAGAGGGAGGCCCATACGAAGGCCACAATACCGTAAAG	135
Db	22	AsnPheValIleGluGlyGluGlyLysGlyLysProPheGluGlyThrGlnThrIleAsn	41
QY	136	CTTAAGGTAAACCAAGGGGGACCTTTGCCATTTCCTTGGGATATTTTGTCCACCACAATTT	195
Db	42	LeuThrValLysGluGlyGlyProLeuProPheAlaTyrAspIleLeuThrAlaAlaPhe	61
QY	196	CAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCAGACTATATAAAAGCTG	255
Db	62	GlnTyrGlyAsnArgAlaPheThrLysTyrProArgAspIleAlaAspTyrPheLysGln	81
QY	256	TCATTTCCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAAGACGGTGGCGTCGTT	315
Db	82	SerPheProGluGlyTyrSerTrpGluArgSerMetThrTyrGluAspGlnGlyIleCys	101
QY	316	ACTGTAACCCAGGATTCAGATTTCAGGATGGCTGTTTCATCTACAAAGTCAAGTTTCATT	375
Db	102	IleIleLysSerAspIleArgMetGluGlyAspCysPheIleTyrGluIleArgTyrAsp	121
QY	376	GGCGTGAACTTTCCTTCCGATGGACCTGTATGCAAAAAGAACACAATGGGCTGGGAAGCC	435
Db	122	GlyValAsnPheProProSerGlyProValMetGlnLysLysThrLeuLysTrpGluPro	141
QY	436	AGCACTGAGCGTTTGTCCTCGTGATGGCGTGTGAAAGGAGAGATTTCATAAGGCTCTG	495
Db	142	SerThrGluLysMetTyrValArgAspGlyValLeuLysGlyGluValAsnMetAlaLeu	161
QY	496	AAGCTGAAAGACGGTGGTCATTACCTAGTTCAATTCAAAAGTATTTACATGGCAAGAAG	555
Db	162	LeuLeuGluGlyGlyGlyHisTyrArgCysAspPheArgSerThrTyrLysAlaLysLys	181
QY	556	CCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGATATAACAAGCCACAAC	615
Db	182	ArgValGlnLeuProAspTyrHisPheValAspHisArgIleGluIleLeuSerHisAsp	201
QY	616	GAAGACTATACAATCGTTTGAGCAGTATGAAGAACCAGGACCGCCACCATCTG	669
Db	202	AsnAspTyrAsnThrValLysLeuSerGluAsnAlaGluAlaArgTyrSerMet	219

Search completed: June 30, 2005, 16:04:10  
Job time : 166.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: June 30, 2005, 15:46:45 ; Search time 136.5 Seconds  
(without alignments)  
3842.103 Million cell updates/sec

Title: US-10-006-922A-11  
Perfect score: 1215  
Sequence: 1 atgaggtcttccaagaatgt.....gccaccatctgttccttaa 678

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US10006922/runat\_30062005\_110551\_9404/app\_query.fasta\_1.839  
-DB=A Geneseq -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10006922@cgn\_1\_1\_224@runat\_30062005\_110551\_9404 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A\_Geneseq\_16Dec04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1214	99.9	225	3 AAY99836	Aay99836 Discosoma
2	1214	99.9	225	3 AAB01622	Aab01622 Discosoma
3	1214	99.9	225	4 AAG65509	Aag65509 Anthozoan
4	1214	99.9	225	5 ABB08834	Abb08834 Yeast opt
5	1214	99.9	225	5 AAE28833	Aae28833 Discosoma
6	1214	99.9	225	5 AAE17540	Aae17540 Discosoma
7	1214	99.9	225	5 AAO18270	Aao18270 Discosoma
8	1214	99.9	225	6 AAE34962	Aae34962 Discosoma
9	1214	99.9	225	7 ADC24126	Adc24126 Discosoma
10	1214	99.9	225	7 ABW00918	Abw00918 Discosoma

11	1214	99.9	225	7 ADF70403	Adf70403 Discosoma
12	1214	99.9	225	7 ADH34489	Adh34489 Discosoma
13	1214	99.9	225	7 ADL46203	Adl46203 Discosoma
14	1214	99.9	225	7 ADN33979	Adn33979 Wild-type
15	1214	99.9	225	8 ADI36421	Adi36421 Discosoma
16	1214	99.9	225	8 ADM97769	Adm97769 D sp red
17	1214	99.9	225	8 ADQ28780	Adq28780 Jellyfish
18	1214	99.9	487	5 ABB08821	Abb08821 Autofluor
19	1214	99.9	506	5 ABB08822	Abb08822 Autofluor
20	1214	99.9	547	5 ABB08823	Abb08823 Autofluor
21	1211	99.7	225	5 AAE28920	Aae28920 Discosoma
22	1211	99.7	225	7 ABW00937	Abw00937 Discosoma
23	1211	99.7	225	7 ABW00929	Abw00929 Discosoma
24	1211	99.7	225	7 ABW00938	Abw00938 Discosoma
25	1211	99.7	225	7 ADH34498	Adh34498 Discosoma
26	1211	99.7	225	7 ADH34499	Adh34499 Discosoma
27	1210	99.6	225	5 AAE28919	Aae28919 Discosoma
28	1210	99.6	225	7 ABW00930	Abw00930 Discosoma
29	1210	99.6	226	4 AAG65510	Aag65510 Anthozoan
30	1210	99.6	242	7 ADE24109	Ade24109 Discosoma
31	1210	99.6	545	7 ADL18132	Adl18132 RFP:PS(NI
32	1210	99.6	548	7 ADL18156	Adl18156 RFP:PS(HI
33	1209	99.5	225	5 AAE28922	Aae28922 Discosoma
34	1209	99.5	225	7 ABW00936	Abw00936 Discosoma
35	1209	99.5	225	7 ABW00931	Abw00931 Discosoma
36	1209	99.5	226	5 ABB08835	Abb08835 Yeast opt
37	1208	99.4	225	7 ABW00932	Abw00932 Discosoma
38	1208	99.4	225	7 ABW00935	Abw00935 Discosoma
39	1208	99.4	225	7 ABW00939	Abw00939 Discosoma
40	1208	99.4	225	7 ADH34500	Adh34500 Discosoma
41	1207	99.3	225	5 AAE28921	Aae28921 Discosoma
42	1207	99.3	225	5 AAE17541	Aae17541 Discosoma
43	1207	99.3	225	7 ABW00941	Abw00941 Discosoma
44	1207	99.3	225	7 ABW00940	Abw00940 Discosoma
45	1207	99.3	225	7 ADL46222	Adl46222 Discosoma

ALIGNMENTS

RESULT 1  
AAY99836  
ID AAY99836 standard; protein; 225 AA.  
XX  
AC AAY99836;  
XX  
DT 12-SEP-2003 (revised)  
DT 19-SEP-2000 (first entry)  
XX  
DE Discosoma sp. "red" novel fluorescent protein drFP583.  
XX  
KW Anthozoa; drFP583; fluorescent protein; non-bioluminescent organism;  
KW fluorescent labeling.  
XX  
OS Discosoma sp; "red".  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 122 /note= "encoded by TC"  
FT Misc-difference 127 /note= "encoded by GTTG"

PI Lukyanoy SA, Fradkov AF, Labas YA, Matz MV, Green G, Chen Y;





US-10-006-922A-11 (1-678) x AAB01622 (1-225)		
QY	1	ATGAGGCTCTTCCAAGATGTTATCAAGGAGTTCATGAGGTTTAAGGTTTCGCATGGAAGGA 60
Db	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
QY	61	ACGGTCAATGGGCACGAGTTTGAATAGAAAGCGAAGGAGAGGGAGGCCATACGAAGGC 120
Db	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
QY	121	CACAATACCGTAAAGCTTAAGGTAAGCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180
Db	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60
QY	181	TTGTCAACCAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240
Db	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80
QY	241	GACTATAAAAGCTGTCTATTTCCCTGAAGGATTTAAATGGGAAGGGTCATGAACCTTTGAA 300
Db	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100
QY	301	GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360
Db	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
QY	361	AAGGTCAAGTTTCATTTGGCGTGAACTTTCCTTCGATGGACCTGTTTGCATTTGGGATATT 180
Db	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
QY	421	ATGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480
Db	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
QY	481	ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTGCATTACCTAGTTGAATTCAAAAGTATT 540
Db	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
QY	541	TACATGGCAAAGAGCCTGTGCAGTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600
Db	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
QY	601	ATAACAAGCCACAACGAAGACTATACAAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC 660
Db	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
QY	661	CACCATCTGTTCCCTT 675
Db	221	HisHisLeuPheLeu 225
RESULT 3		
AAG65509		
ID	AAG65509	standard; protein; 225 AA.
XX		
AC	AAG65509;	
XX		
DT	30-NOV-2001	(first entry)
XX		
DE	Anthozoan red fluorescent protein sequence.	
XX		
KW	Fluorescent protein; Anthozoan; fluorescence; marker; FRET; red.	
XX		
OS	Anthozoa.	
XX		
PN	WO200162919-A1.	
XX		
PD	30-AUG-2001.	
XX		
PF	13-FEB-2001; 2001WO-US004625.	
XX		
PR	23-FEB-2000; 2000US-0184732P.	
XX		

FA	(AURO-) AURORA BIOSCIENCES CORP.	
XX		
PI	Nelson D, Zamaira E, Tsien R;	
XX		
DR	WPI; 2001-557704/62.	
XX		
PT	Proteins for Fluorescence Resonance Energy Transfer (FRET) comprise functional red fluorescent proteins, and the encoding nucleic acids, with key mutations for improving the proteins function.	
PT		
XX		
PS	Disclosure; Page 85; 90pp; English.	
XX		
CC	The invention provides a nucleic acid encoding functional red fluorescent CC protein (II) that differs from the sequence of an Anthozoan red fluorescent protein by at least one amino acid substitution, and with different fluorescent properties. The red fluorescent protein of the invention can be expressed by standard recombinant methodology. (II) are used a fluorescent markers and FRET partners. It is used for identifying protein-protein interactions. (II) is also suitable for multiplexed fluorescent analysis and FRET-based applications using existing Aequorea fluorescent proteins. (II) has improved brightness, reduced spectral cross talk, and is rapidly and efficiently expressed in mammalian cells. The key mutations in the encoding nucleic acids provide improved folding, brightness, and create (II) with sharper, more defined excitation and emission peaks when expressed in mammalian cells. The present sequence represents an anthozoan fluorescent protein	
XX		
SQ	Sequence 225 AA;	
Alignment Scores:		
Pred. No.:	3.01e-135	
Score:	1214.00	
Percent Similarity:	100.00%	
Best Local Similarity:	100.00%	
Query Match:	99.92%	
DB:	4	
US-10-006-922A-11 (1-678) x AAG65509 (1-225)		
QY	1	ATGAGGCTCTTCCAAGATGTTATCAAGGAGTTCATGAGGTTTAAGGTTTCGCATGGAAGGA 60
Db	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
QY	61	ACGGTCAATGGGCACGAGTTTGAATAGAAAGCGAAGGAGAGGGGAGGCCATACGAAGGC 120
Db	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
QY	121	CACAATACCGTAAAGCTTAAGGTAAGCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180
Db	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60
QY	181	TTGTCAACCAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240
Db	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80
QY	241	GACTATAAAAGCTGTCTATTTCCCTGAAGGATTTAAATGGGAAGGGTCATGAACCTTTGAA 300
Db	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100
QY	301	GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360
Db	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
QY	361	AAGGTCAAGTTTCATTTGGCGTGAACTTTCCTTCGATGGACCTGTTTGCATTTGGGATATT 180
Db	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
QY	421	ATGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480
Db	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
QY	481	ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTGCATTACCTAGTTGAATTCAAAAGTATT 540
Db	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
QY	541	TACATGGCAAAGAGCCTGTGCAGTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600
Db	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
QY	601	ATAACAAGCCACAACGAAGACTATACAAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC 660
Db	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
QY	661	CACCATCTGTTCCCTT 675
Db	221	HisHisLeuPheLeu 225

Db 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
QY 541 TACATGGCAAAGAGCTGTGCAGCTACCAAGGGTACTACTATGTTGACTCCAAACTGGAT 600  
Db 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrValAspSerLysLeuAsp 200  
QY 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC 660  
Db 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
QY 661 CACCATCTGTTCCCTT 675  
Db 221 HisHisLeuPheLeu 225

RESULT 4  
ABB08834  
ID ABB08834 standard; protein; 225 AA.  
AC ABB08834;

DT 29-MAY-2002 (first entry)

XX Yeast optimised RFP SEQ ID NO 17.

KW Yeast; red fluorescent protein; RFP; plant; transgenic; GFP;  
KW Saccharomyces cerevisiae; Nicotiana tabacum; Arabidopsis thaliana;  
KW Escherichia coli; green fluorescent protein; biotechnology.

XX Anthozoa.

OS DE20001395-U1.

PN 15-MAR-2001.

XX 27-JAN-2000; 2000DE-02001395.

XX 27-JAN-2000; 2000DE-02001395.

PR (GPCB-) GPC BIOTECH AG.

XX WPI; 2002-228394/29.

DR N-PSDB; ABA95905, ABA95921, ABA95922.

XX New DNA encoding red fluorescent protein, useful as marker in  
PT biotechnology, has sequence optimized for expression in eukaryotes,  
PT especially yeast or plants.

XX Disclosure; Page 13-14; 19pp; German.

XX The invention relates to DNA (I) containing either sequence ABA95905 or  
CC sequence ABA95906 encoding a yeast optimised Red Fluorescent Protein  
CC (yRFP). (I) are used to express red fluorescent protein (RFP) in  
CC eukaryotes, especially yeast, especially Saccharomyces cerevisiae and  
CC plants, especially dicotyledonous plants including Nicotiana tabacum or  
CC Arabidopsis thaliana and also in prokaryotes, especially bacteria,  
CC especially Escherichia coli. RFP is useful in the same way as green  
CC fluorescent protein but is more generally applicable in modern  
CC biotechnology. (I) are optimised for expression in yeast and so generate  
CC RFP at higher levels with stronger fluorescence and thus lowers the  
CC detection limit and gives a better signal-to-noise ratio. The present  
CC sequence is that of the yeast optimised RFP

XX Sequence 225 AA;

Alignment Scores:  
Pred. No.: 3.01e-135 Length: 225  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.92% Indels: 0  
DB: 5 Gaps: 0

US-10-006-922A-11 (1-678) x ABB08834 (1-225)

QY 1 ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTTCATGAGGTTTAAAGTTTCGCATGGAAGGA 60  
Db 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
QY 61 ACGGTCAATGGGCACGAGTTTGAATAGAAAGCGGAAGGAGAGGGAGGCCATACGAAGGC 120  
Db 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
QY 121 CACAATACCGTAAAGCTTAAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
Db 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
QY 181 TTGTCAACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA 240  
Db 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
QY 241 GACTATAAAAAAGCTGTCTATTTCCCTGAAGGATTTAAATGGGAAAAGGGTCATGAAC TTGAA 300  
Db 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
QY 301 GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTTGCAGGATGGCTGTTTCATCTAC 360  
Db 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
QY 361 AAGGTCAGTTCATTGGCGTGAACTTCCCTTCCGATGGACCTGTTATGCAAAAGAAGACA 420  
Db 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGATCCTCGTATGGCGTGTGAAAGGAGAG 480  
Db 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
QY 481 ATTCAATAAGSCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT 540  
Db 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
QY 541 TACATGGCAAAAGACCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAT 600  
Db 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrValAspSerLysLeuAsp 200  
QY 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC 660  
Db 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
QY 661 CACCATCTGTTCCCTT 675  
Db 221 HisHisLeuPheLeu 225

RESULT 5  
AAE28833  
ID AAE28833 standard; protein; 225 AA.

XX AAE28833;

XX 27-DEC-2002 (first entry)

DE Discosoma sp. drFP583 (NFP-6) wild-type protein.

XX Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
KW fluorescence activated cell sorting application; fluorescent timer;  
KW biosensor; fluorescence resonance energy transfer application; FRET;  
KW colouring agent; recombinant DNA application; analyte detection assay;  
KW sunscreen; second messenger detector; drFP583 protein; NFP-6.

XX Discosoma sp.

XX WO200268459-A2.

XX 06-SEP-2002.

XX 20-FEB-2002; 2002WO-US005749.

XX



PR 21-FEB-2001; 2001US-0270983P.  
PR 04-DEC-2001; 2001US-00006922.  
XX  
PA (CLON-) CLONTECH LAB INC.  
XX  
PI Lukyanov S, Lukyanov K, Yanushevich Y, Savitsky A, Fradkov A;  
XX  
DR WPI; 2002-691654/74.  
DR N-PSDB; AAD46278.  
XX  
PT New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
PT of an aggregating Cnidarian chromo- or fluorescent protein or mutant for  
PT analyte detection assays or fluorescence activated cell sorting  
PT applications;  
XX  
PS Disclosure; Page 70-71; 80pp; English.  
XX  
CC The invention relates to nucleic acid molecules encoding non-aggregating  
CC chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
CC useful in analyte detection assays, as colouring agents, as markers in  
CC recombinant DNA applications, as sunscreens or filters, in fluorescence  
CC resonance energy transfer (FRET) applications, as biosensors in  
CC prokaryotic and eukaryotic cells, in screening assays, as second  
CC messenger detectors, in fluorescence activated cell sorting applications,  
CC in protease cleavage assays or as fluorescent timers. The present  
CC sequence is Discosoma sp. drFP583 (NFP-6) wild-type protein of the  
CC invention  
XX  
SQ Sequence 225 AA;  
  
Alignment Scores:  
Pred. No.: 3.01e-135 Length: 225  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.92% Indels: 0  
DB: 5 Gaps: 0  
  
US-10-006-922A-11 (1-678) x AAE28833 (1-225)  
QY 1 ATGAGGTCTTCCAAGATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGTCATGAAGGA 60  
DB 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
QY 61 ACGGTCAATGGGACGAGTTTGAATAGAGCGAAGGAGAGGGAGGCCATACGAAGGC 120  
DB 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
QY 121 CACAATACCGTAAAGCTTAAGGTAAACCAAGGGGGACCTTTGCCATTTGCTGGATATT 180  
DB 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspile 60  
QY 181 TTGTCAACCAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCTTCGCGACATACCA 240  
DB 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspilePro 80  
QY 241 GACTATAAAAGCTGTCATTTCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTGAA 300  
DB 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
QY 301 GACGGTGGCGTCTACTGTAAACCCAGGATTCAGTTTCAGGATGGCTGTTTCATCTAC 360  
DB 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
QY 361 AAGGTCAAGTTTCATGGCGTGAACCTTTCCTTCGATGGACCTGTTATGCAAAAGAAGACA 420  
DB 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
QY 421 ATGGGCTGGGAAGCCACACTGAGCGTTTGTATCTCGTATGGCGTGTGAAAGGAGAG 480  
DB 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
QY 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTGTCATTACCTAGTTGAATCAAAAGTATT 540

Db 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
QY 541 TACATGGCAAGAAGCCTGTGCAGTACCAGGTACTACTATGTGACTCCAAACTGGAT 600  
Db 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrValAspSerLysLeuAsp 200  
QY 601 ATAACAAGCCACAACGAAGACTATACAAATCGTTGAGCAGTATGAAGAACCGAGGACGC 660  
Db 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
QY 661 CACCATCTGTTCTCTT 675  
Db 221 HisHisLeuPheLeu 225  
  
RESULT 6  
AAE17540  
ID AAE17540 standard; protein; 225 AA.  
XX  
AC AAE17540;  
XX  
DT 22-APR-2002 (first entry)  
XX  
DE Discosoma sp. humanised wild-type Anthozoa protein drFP583.  
XX  
KW Fluorescent timer protein; protein movement; translocation; trafficking;  
KW promoter activity; gene expression; transgenic plant; gene modification;  
KW protein age; anthozoa protein; drFP583.  
XX  
OS Discosoma sp.  
XX  
PN WO200196373-A2.  
XX  
PD 20-DEC-2001.  
XX  
PF 13-JUN-2001; 2001WO-US019097.  
XX  
PR 14-JUN-2000; 2000US-0211607P.  
XX  
PA (CLON-) CLONTECH LAB INC.  
XX  
PI Pradkov AF, Terskikh A;  
XX  
DR WPI; 2002-154595/20.  
DR N-PSDB; AAD28207.  
XX  
PT New fluorescent timer proteins comprising an emission spectrum that  
PT changes over time from a first wavelength to a second wavelength, useful  
PT for monitoring intracellular protein movement, translocation, trafficking  
PT or stability.  
XX  
PS Example 1; Fig 1; 89pp; English.  
XX  
CC The invention relates to a fluorescent timer protein having an emission  
CC spectrum that changes over time after synthesis from a first wavelength  
CC to a second wavelength. The fluorescent timer proteins are useful in  
CC monitoring the activity of a promoter, determining the age of a protein,  
CC identifying an agent that modulates the activity of a promoter and in  
CC enriching a population of cells comprising a fluorescent timer protein.  
CC The fluorescent timer proteins are also useful for assessing gene  
CC expression during development of a multicellular organism or during  
CC cellular differentiation, in response to a drug or other inducer of  
CC promoter activity, as a reporter to serve as a read-out of promoter  
CC activity, monitoring intracellular protein movement or translocation,  
CC protein trafficking, or protein stability, to investigate temporal  
CC aspects of the activity of a regulatory element, for determining cell  
CC fate during development and organ remodelling, in spatial and temporal  
CC visualisation of newly synthesised proteins and accumulated proteins, and  
CC in distinguishing between newly formed and pre-existing structures, e.g.  
CC membrane junctions and extracellular matrix components. The fluorescent  
CC timer proteins may further be used to investigations where photobleaching  
CC techniques are employed, as detectable labels, as selectable markers, as  
CC biosensors in prokaryotic and eukaryotic cells, in protease cleavage



CC assays, and as second messenger detectors. The nucleic acids can be used  
CC to generate transgenic, non-human plants or animals or site-specific gene  
CC modifications in cell lines. The present sequence is Discosoma sp.  
CC humanised wild-type Anthozoa protein drFP583 used for generating  
CC fluorescent proteins  
XX  
SQ Sequence 225 AA;

Alignment Scores:  
Pred. No.: 3.01e-135 Length: 225  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.92% Indels: 0  
DB: 5 Gaps: 0

US-10-006-922A-11 (1-678) x AAE17540 (1-225)

QY 1 ATGAGGTCTTCCAAGATGTTATCAAGGAGTTCATGAGGTTTAAGGTTTCGCATGGAAGGA 60  
Db 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
QY 61 ACGGTCAATGGGCACGAGTTTGAATAGAGCGAAGGAGAGGGAGGCCATACGAAGGC 120  
Db 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTTGGGTATTT 180  
Db 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
QY 181 TTGTCACCACAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240  
Db 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
QY 241 GACTATAAAAAGCTGTCATTTCTGAAGGATTTAAATGGGAAAGGTCATGAACTTTGAA 300  
Db 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
QY 301 GACGGTGGCGTCGTTACTGTAACCCAGGATTCAGTTGCAGGATGGCTGTTTCATCTAC 360  
Db 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
QY 361 AAGGTCAAAGTTTCATGGCGTGAACTTTCCTCCGATGGACCTGTTATGCAAAAAGAGACA 420  
Db 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
Db 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
QY 481 ATTCAATAGGCTCTGAAGCTGAAAGACGGTGGTCACTTACCTAGTTGAATTCAAAAGTATT 540  
Db 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
QY 541 TACATGGCAAAGACCTGTGTCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600  
Db 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
QY 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC 660  
Db 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
QY 661 CACCATCTGTTTCCTT 675  
Db 221 HisHisLeuPheLeu 225

RESULT 7  
AAO18270  
ID AAO18270 standard; protein; 225 AA.

XX  
AC AAO18270;  
XX  
DT 26-SEP-2002 (first entry)

XX Discosoma red fluorescent protein.  
DE Yeast; RAD54; promoter; genotoxicity cassette; cytotoxicity cassette;  
XX modified yeast strain; environmental pollution.  
KW Discosoma sp.  
OS DE10061872-A1.  
XX 20-JUN-2002.  
PD 12-DEC-2000; 2000DE-01061872.  
XX 12-DEC-2000; 2000DE-01061872.  
PR (LICH/) LICHTENBERG-FRATE H.  
XX Lichtenberg-Frate H;  
PI WPI; 2002-539633/58.  
XX N-PSDB; AAL47952.  
DR Modified yeast strain, useful for detecting toxic compounds in  
PT environment, contains integrated cassettes responsive to genotoxic and  
PT cytotoxic compounds.  
XX Disclosure; Page 21-22; 34pp; German.  
PS The present invention relates to a modified yeast strain that contains,  
XX integrated stably and functionally in its genome, a genotoxicity cassette  
CC and a cytotoxicity cassette, each comprising a promoter and reporter  
CC gene, both of which are different in the two cassettes. The modified  
CC yeast strain is used to detect environmental pollution, especially  
CC genotoxic and/or cytotoxic substances in complex environmental  
CC contaminants, especially organic compounds, but also (non-)ionising  
CC radiation and chemical carcinogens. Particular applications are in  
CC monitoring (waste) water (e.g. as an early warning system), medical  
CC toxicology screening and for industrial process control. The present  
CC sequence is a marker protein suitable for use in the cassettes of the  
CC present invention  
XX Sequence 225 AA;

Alignment Scores:  
Pred. No.: 3.01e-135 Length: 225  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.92% Indels: 0  
DB: 5 Gaps: 0

US-10-006-922A-11 (1-678) x AAO18270 (1-225)

QY 1 ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTCATGAGTTTAAGGTTTCGCATGGAAGGA 60  
Db 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
QY 61 ACGGTCAATGGGCACGAGTTTGAATAGAGCGAAGGAGAGGGAGGCCATACGAAGGC 120  
Db 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTTGGGTATTT 180  
Db 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
QY 181 TTGTCACCACAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240  
Db 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
QY 241 GACTATAAAAAGCTGTCATTTCTCTGAAGGATTTAAATGGGAAAGGTCATGAACTTTGAA 300  
Db 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100

QY 301 GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGGTTTGCAGGATGGCTGTTTCATCTAC 360  
Db 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
QY 361 AAGGTCAAGTTCAATGGCGTGAACCTTTCCTTCGATGGACCTGTTATGCAAAAGAGACA 420  
Db 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGTGATGGCGTGTGAAAGGAGAG 480  
Db 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
QY 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAAGTATT 540  
Db 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
QY 541 TACATGGCAAGAGCCTGTGCAGCTACCAGGCTACTACTATGTTGACTCCAAACTGGAT 600  
Db 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
QY 601 ATAACAAGCCACACGAAGACTATACAATCGTTGACGAGTATGAAAGAACCGAGGACGC 660  
Db 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
QY 661 CACCATCTGTTCCCTT 675  
Db 221 HisHisLeuPheLeu 225

RESULT 8  
AAE34962

ID AAE34962 standard; protein; 225 AA.

AC AAE34962;

DT 28-MAY-2003 (first entry)

DE Discosoma species red fluorescent protein (RFP).

KW Phosphorylation indicator; fluorescent protein; detection; phosphatase;  
KW kinase; red fluorescent protein; RFP.

OS Discosoma sp.

PN WO200295058-A2.

PD 28-NOV-2002.

PF 24-MAY-2002; 2002WO-US016955.

PR 24-MAY-2001; 2001US-00865291.

PA (REGC ) UNIV CALIFORNIA.

PI Tsien RY, Ting AY, Zhang J;

DR WPI; 2003-148474/14.

DR N-PSDB; AAD53432.

PT Novel chimeric phosphorylation indicators, useful for detecting  
PT kinase/phosphatase in samples, has donor molecule, phosphorylatable  
PT domain, phosphoaminoacid binding domain, and acceptor molecule, in  
PT operative linkage.

PS Disclosure; Col 65-66; 38pp; English.

CC The present invention relates to chimeric phosphorylation indicators  
CC comprising a phosphorylation polypeptide and a fluorescent protein or in  
CC operative linkage, a donor molecule, a phosphorylatable domain, a  
CC phosphoaminoacid binding domain (PAABD) and an acceptor molecule. The  
CC phosphorylation indicators of the invention are useful for detecting  
CC kinases or phosphatases in a biological sample. They are also useful in  
CC high throughput analysis e.g. for detecting a kinase inhibitor or

CC phosphatase inhibitor. The present sequence is Discosoma species red  
CC fluorescent protein (RFP) used in the invention  
XX  
SQ Sequence 225 AA;

Alignment Scores:  
Pred. No.: 3,01e-135 Length: 225  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.92% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-11 (1-678) x AAE34962 (1-225)

QY 1 ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTCCGATGGAAGGA 60

Db 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20

QY 61 ACGGTCAATGGGCACGAGTTTGAATAGAACGCGAAGGAGGGGAGGCCATACGAAGGC 120

Db 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40

QY 121 CACAATACCGTAAAGCTTAAGGTAAACCAAGGGGGGACCTTTGCCATTGCTGGGATATT 180

Db 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60

QY 181 TTGTCAACCAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240

Db 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80

QY 241 GACTATAAAAAGCTGTTCATTTTCCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAA 300

Db 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100

QY 301 GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360

Db 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120

QY 361 AAGGTCAAGTTTCATTTGGCGTGAACCTTTCCTCCGATGGACCTGTTATGCAAAAGAGACA 420

Db 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140

QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGTGCGTGTGAAAGGAGAG 480

Db 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160

QY 481 ATTCATAAGSCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAAGTATT 540

Db 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180

QY 541 TACATGGCAAAAGCCCTGTGCAGCTACCAGGTTACTACTATGTTGACTCCAAACTGGAT 600

Db 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200

QY 601 ATAACAAGCCACAAACGAAGACTATACAATCGTTGAGCAGTATGAAAGACCGAGGACGC 660

Db 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220

QY 661 CACCATCTGTTCCCTT 675

Db 221 HisHisLeuPheLeu 225

RESULT 9  
ADC24126

ID ADC24126 standard; protein; 225 AA.

XX ADC24126;

AC ADC24126;

XX 18-DEC-2003 (first entry)

DE Discosoma wild-type red fluorescent protein.

XX

KW Discosoma red fluorescent protein; DsRed; AB interface; AC interface;  
KW fluorescent protein variant; transcription induction detection;  
KW fluorescence energy resonance transfer; FRET; protein kinase;  
XX protein phosphatase; ion indicator.  
OS Discosoma.  
XX  
PN US2003059835-A1.  
XX  
PD 27-MAR-2003.  
XX  
PF 10-APR-2002; 2002US-00121258.  
XX  
PR 26-FEB-2001; 2001US-00794308.  
PR 24-MAY-2001; 2001US-00866538.  
XX  
PA (TSIE/) TSIENT R Y.  
PA (CAMP/) CAMPBELL R E.  
XX  
PI Tsien RY, Campbell RE;  
XX  
DR WPI; 2003-743764/70.  
DR N-PSDB; ADC24127, ADC24134.  
XX  
PT Novel polynucleotide sequence encoding Discosoma red fluorescent protein  
PT variant having a reduced propensity to oligomerize, useful for detecting  
PT transcriptional activity.  
XX  
PS Claim 1; SEQ ID NO 1; 67pp; English.  
XX  
CC The invention describes a polynucleotide sequence (I) encoding a  
CC Discosoma red fluorescent protein (DsRed) variant having a reduced  
CC propensity to oligomerise, comprising amino acid substitutions at the AB  
CC and/or AC interfaces of the wild-type DsRed sequence (S1) comprising 225  
CC amino acids, given in the specification, where the substitutions result  
CC in reduced propensity of the DsRed variant to form tetramers. (I) is  
CC useful for detecting transcriptional activity by providing a host cells  
CC conatinig a vector which comprises (I) operatively linked to an  
CC expression control sequence, and an unit to assay the variant fluorescent  
CC protein fluorescence, and assaying fluorescence of the variant  
CC fluorescent protein produced by (VII), where variant fluorescent protein  
CC fluorescence is indicative of transcriptional activity. A polynucleotide  
CC encoding a fusion protein is useful for the analysis of in vivo  
CC localisation or trafficking of a polypeptide of interest. A polypeptide  
CC marker is useful as markers to identify the location and amount of a  
CC target protein produced, where the target protein is fused to the marker,  
CC as a complement to or alternative for the green fluorescent protein or  
CC its spectral variant, for detecting induction of transcriptions, in  
CC applications involving fluorescence energy resonance transfer (FRET),  
CC which detects events as the function of the movement of fluorescent  
CC donors and acceptors towards or away from each other, for making  
CC fluorescent sensors for protein kinase and phosphatase activities or  
CC indicators for ions and molecules such as Ca2+, Zn2+, for identifying the  
CC presence of a molecule in a sample, for identifying a specific  
CC interaction of a first and second molecule, for determining whether a  
CC sample contains an enzyme or for determining the pH of the sample. (I) is  
CC useful for identifying a region or condition that regulates the activity  
CC of an expression control sequence. This is the amino acid sequence of  
CC Discosoma wild-type red fluorescent protein.  
XX  
SQ Sequence 225 AA;

Alignment Scores:  
Pred. No.: 3.01e-135 Length: 225  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.92% Indels: 0  
DB: 7 Gaps: 0

US-10-006-922A-11 (1-678) x ADC24126 (1-225)

QY 1 ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGATGGAAGGA 60

Db 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
QY 61 ACGGTCATGGGCACGAGTTTGAATAGAACCGGAAGGAGAGGGGAGCCATACGAAGGC 120  
Db 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTGCTTGGGATATT 180  
Db 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
QY 181 TTGTCACCAACAATTTCAGTATGGAAAGCAAGGTATATGTCAAGCACCTGCCGACATACCA 240  
Db 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
QY 241 GACTATAAAAGCTGTCAATTCCCTGAAGGATTAAATGGGAAGGTCATGAACCTTTTGAA 300  
Db 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
QY 301 GACGGTGGCGTCGTTACTGTAAACCCAGGATTCCAGTTTGCAGGATGGCTGTTTCATCTAC 360  
Db 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
QY 361 AAGGTCAAGTTCATTGGCGTGAACCTTCCTCCGATGGACCTGTATGCAAAAGAAGACA 420  
Db 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
Db 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
QY 481 ATTCAATAAGGCTCTGAAGCTGAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT 540  
Db 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
QY 541 TACATGGCAAAAGACCTGTGCAGCTACCAGGTACTACTATGTGACTCCAAACTGGAT 600  
Db 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
QY 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCAGGGACGC 660  
Db 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
QY 661 CACCATCTGTTCCTTT 675  
Db 221 HisHisLeuPheLeu 225

RESULT 10  
ABW00918  
ID ABW00918 standard; protein; 225 AA.  
XX  
AC ABW00918;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE Discosoma sp. red fluorescent protein (RFP).  
XX  
KW Fluorescent protein; resonance energy transfer; pH; detection;  
KW red fluorescent protein; RFP.  
XX  
OS Discosoma sp.  
XX  
PN US2003170911-A1.  
XX  
PD 11-SEP-2003.  
XX  
PF 26-FEB-2001; 2001US-00794308.  
XX  
PR 26-FEB-2001; 2001US-00794308.  
XX  
PA (TSIE/) TSIENT R Y.  
PA (ZACH/) ZACHARIAS D A.  
PA (BAIR/) BAIRD G S.



XX Tsien RY, Zacharias DA, Baird GS;  
PI WPI; 2003-802418/75.  
XX N-PSDB; AAD61969.  
PT Fluorescent proteins containing a mutation that reduces or eliminates its  
PT ability to oligomerize which gives more reliable fluorescence resonance  
PT energy transfer results and are useful to detect molecule interaction,  
PT enzymes, or sample pH.  
XX Claim 10; Page 30-31; Opp; English.  
PS The invention relates to a non-oligomerising fluorescent protein  
XX containing a mutation that reduces or eliminates its ability to  
CC oligomerise. The fluorescent protein gives more reliable fluorescence  
CC resonance energy transfer results and are useful to detect molecule  
CC interaction, enzymes, or sample pH. These are also used to identify  
CC agents or conditions that regulate expression of control sequences. The  
CC present sequence is Discosoma sp. red fluorescent protein (RFP)  
XX  
SQ Sequence 225 AA;  
  
Alignment Scores:  
Pred. No.: 3.01e-135 Length: 225  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.92% Indels: 0  
DB: 7 Gaps: 0  
  
US-10-006-922A-11 (1-678) x ABW00918 (1-225)  
  
QY 1 ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTTCATGAGTTTAAGGTTTCGCATGGAAGGA 60  
Db 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
  
QY 61 ACGGTCAATGGCAGCAGTTTGAAATAGAGCGAAGGAGAGAGGGGAGGCCATACGAAGGC 120  
Db 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
  
QY 121 CACAATACCGTAAAGCTTAAGTTAACCAAGGGGGACCTTTGCCATTGCTTGGGATATT 180  
Db 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
  
QY 181 TTGTCACCACTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCTGCCGACATACCA 240  
Db 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
  
QY 241 GACTATAAAAGCTGTCTATTCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTGAA 300  
Db 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
  
QY 301 GACGGTGGCGTCTGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360  
Db 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
  
QY 361 AAGGTCAAGTTTCATGGCGTGAACCTTTCCTCCGATGGACCTGTTATGCAAAAGAAGACA 420  
Db 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
  
QY 421 ATGGCTGGGAGCAGCACTGAGCGTTTGTATCCTCGTATGCGGTGTGAAAGGAGAG 480  
Db 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
  
QY 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTTCATTACCTAGTTTGAATTCAAAAGTATT 540  
Db 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
  
QY 541 TACATGGCAAAGAACCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAT 600  
Db 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200

QY 601 ATAACAAGCCACAACGAAGACTATACAAATCGTTGAGCAGTATGAAGAACCAGGAGCGC 660  
Db 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
  
QY 661 CACCATCTGTTCTCTT 675  
Db 221 HisHisLeuPheLeu 225  
  
RESULT 11  
ADF70403  
ID ADF70403 standard; protein; 225 AA.  
XX  
AC ADF70403;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE Discosoma wild-type GFP variant protein SeqID26.  
XX  
KW ligand; orphan receptor protein; fusion protein; fluorescent protein;  
KW cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;  
KW GFPuv; Enhanced GFP; EGFP.  
XX  
OS Discosoma sp.  
XX  
PN WO2003071272-A1.  
XX  
PD 28-AUG-2003.  
XX  
PF 21-FEB-2003; 2003WO-JP001901.  
XX  
PR 22-FEB-2002; 2002JP-00045728.  
PR 23-JUL-2002; 2002JP-00213949.  
PR 11-OCT-2002; 2002JP-00298237.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Hinuma S, Fujii R, Ogi K, Komatsu H, Kawamata Y, Hoshiya M;  
XX WPI; 2003-697654/66.  
DR N-PSDB; ADF70404.  
XX  
PT Transformation of cells with a fusion protein of an orphan receptor  
PT protein with a fluorescent protein useful for identification of ligands  
PT to the orphan receptor.  
XX  
PS Disclosure; SEQ ID NO 26; 594pp; Japanese.  
XX  
CC This invention relates to a novel method of identifying ligands to an  
CC orphan receptor protein which comprises transforming cells with DNA  
CC encoding a fusion protein of the orphan receptor with a fluorescent  
CC protein, so that the fusion protein is expressed in the cells (or cell  
CC membranes isolated from them) and contacting the cells with the potential  
CC ligand to be tested. A suitable fluorescent protein for incorporation in  
CC the fusion protein is green fluorescent protein (GFP), for example GFP-1,  
CC wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the  
CC identification of ligands binding to an orphan receptor protein.  
XX  
SQ Sequence 225 AA;  
  
Alignment Scores:  
Pred. No.: 3.01e-135 Length: 225  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.92% Indels: 0  
DB: 7 Gaps: 0  
  
US-10-006-922A-11 (1-678) x ADF70403 (1-225)  
  
QY 1 ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTTCATGAGTTTAAGGTTTCGCATGGAAGGA 60  
Db 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20





Query Match:	99.92%	Indels:	0
DB:	7	Gaps:	0
US-10-006-922A-11 (1-678) x ADH34489 (1-225)			
QY	1	ATGAGGTCTTCCAAAGATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGATGGAAGGA	60
Db	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
QY	61	ACGGTCAATGGCACGAGTTTGAAATAGAAAGGCAAGGAGAGGGGAGGCCATACGAAGGC	120
Db	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
QY	121	CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT	180
Db	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle	60
QY	181	TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA	240
Db	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
QY	241	GACTATAAAAGCTGTCATTTCTGAAGGATTTCAAGGATTTGAAGGTTTCATGAACCTTTGAA	300
Db	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
QY	301	GACGGTGGCGTCTACTGTACTGTAAACCCAGGATTCAGGATTTGCAGGATGGCTGTTTCATCTAC	360
Db	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
QY	361	AAGGTCAAGTTCATTGGCGTGAACTTTCCCTTCCGATGGACCTGTTATGCAAAAGAACACA	420
Db	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
QY	421	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG	480
Db	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
QY	481	ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTTCATTACCTAGTTGAATCAAAAGTATT	540
Db	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
QY	541	TACATGGCAAAGAGCCTGTGCAGCTACCAAGGTPACTACTATGTTGACTCCAAAACCTGGAT	600
Db	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
QY	601	ATAACAAGCCACAACGAAGACTATACAATCGTTGACGAGTATGAAAGAACCGGAGCGC	660
Db	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
QY	661	CACCATCTGTTCCIT 675	
Db	221	HisHisLeuPheLeu 225	
RESULT 13			
ID	ADL46203	standard; protein; 225 AA.	
XX	AC	ADL46203;	
XX	DT	20-MAY-2004 (first entry)	
XX	DE	Discosoma red fluorescent (DsRed) protein.	
KW	red fluorescent protein; DsRed; fluorescence; red wavelength;		
KW	oligomerization; tetramerization; immunoassay; hybridization assay.		
OS	Discosoma sp.		
XX	PN	WO2003086446-A1.	
XX	PD	23-OCT-2003.	
XX	PF	09-APR-2003; 2003WO-US010879.	

XX	10-APR-2002; 2002US-00121258.		
PR	29-JUL-2002; 2002US-00209208.		
XX	(REGC ) UNIV CALIFORNIA.		
PA	Tsien RY, Campbell RE, Baird GS;		
XX	WPI; 2003-845265/78.		
DR	N-PSDB; ADL46204.		
XX	New monomeric and dimeric Anthozoan fluorescent protein variants with reduced propensity to oligomerize, and encoding polynucleotides, useful in molecular biology, e.g. in immunoassays or in tracking protein movement in cells.		
PS	Claim 1; SEQ ID NO 1; 166pp; English.		
XX	The invention relates to a polynucleotide sequence encoding a Discosoma red fluorescent protein (DsRed) variant having a reduced propensity to oligomerize. The protein variant comprises one or more amino acid substitutions at the AB and/or AC interface(s) of the wild-type DsRed sequence, where the substitutions result in reduced propensity of the DsRed variant to form tetramers and where the variant displays detectable fluorescence of at least one red wavelength. The composition and methods are useful in producing red fluorescent proteins having reduced propensity for oligomerization, especially tetramerization. The protein may be used in molecular biology and in other scientific applications, such as in immunoassays or hybridization assays, or in tracking the movement of proteins in cells. This sequence corresponds to the DsRed protein.		
XX	Sequence 225 AA;		
SQ	Alignment Scores:		
	Pred. No.:	3.01e-135	Length: 225
	Score:	1214.00	Matches: 225
	Percent Similarity:	100.00%	Conservative: 0
	Best Local Similarity:	100.00%	Mismatches: 0
	Query Match:	99.92%	Indels: 0
	DB:	7	Gaps: 0
US-10-006-922A-11 (1-678) x ADL46203 (1-225)			
QY	1	ATGAGGTCTTCCAAAGAATGTTATCAAGGAGTTTCATGAGGTTTAAGTTCGATGGAAGGA	60
Db	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
QY	61	ACGGTCAATGGGCACGAGTTTGAATAGAAAGGCGAAGGAGAGGGGAGGCCATACGAAGGC	120
Db	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
QY	121	CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT	180
Db	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle	60
QY	181	TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA	240
Db	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
QY	241	GACTATAAAAGCTGTCATTTCTGAAGGATTTAAATGGGAAAGGGTCAATGAACCTTTGAA	300
Db	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
QY	301	GACGGTGGCGTCTACTGTACTGTAAACCCAGGATTCAGGATTTGCAGGATGGCTGTTTCATCTAC	360
Db	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
QY	361	AAGGTCAAGTTCATTGGCGTGAACTTTCCCTTCCGATGGACCTGTTATGCAAAAGAACACA	420
Db	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
QY	421	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG	480
Db	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
QY	481	ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTTCATTACCTAGTTGAATCAAAAGTATT	540
Db	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
QY	541	TACATGGCAAAGAGCCTGTGCAGCTACCAAGGTPACTACTATGTTGACTCCAAAACCTGGAT	600
Db	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
QY	601	ATAACAAGCCACAACGAAGACTATACAATCGTTGACGAGTATGAAAGAACCGGAGCGC	660
Db	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
QY	661	CACCATCTGTTCCIT 675	
Db	221	HisHisLeuPheLeu 225	

Db 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
QY 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT 540  
Db 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
QY 541 TACATGGCAAAGACCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAT 600  
Db 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
QY 601 ATAACAAGCCACACGAAGACTATACAATCGTTGAGCAGTATGAAAGACCGAGGGACGC 660  
Db 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
QY 661 CACCATCTGTTCCCTT 675  
Db 221 HisHisLeuPheLeu 225

RESULT 14  
ADN33979  
ID ADN33979 standard; protein; 225 AA.  
XX  
AC ADN33979;  
XX  
DT 01-JUL-2004 (first entry)  
XX  
DE Wild-type DsRED protein.  
XX  
KW Cnidarian; fluorescence resonance energy transfer; FRET; wild-type DsRED.  
XX  
OS Discosoma sp.  
XX  
PN WO2003054158-A2.  
XX  
PD 03-JUL-2003.  
XX  
PF 18-DEC-2002; 2002WO-US040539.  
XX  
PR 19-DEC-2001; 2001US-0341723P.  
XX  
PA (UYCH-) UNIV CHICAGO.  
XX  
PI Bevis B, Glick B;  
XX  
DR WPI; 2003-569236/53.  
DR N-PSDB; ADN33978.  
XX  
PT Novel nucleic acid encoding a rapidly maturing chromo- or fluorescent  
PT mutant of a Cnidarian chromo- or fluorescent protein or its mutant,  
PT useful for applications involving chromo- or fluorescent proteins.  
XX  
PS Claim 8; SEQ ID NO 2; 65pp; English.

CC The present invention relates to nucleic acid that encodes a rapidly  
CC maturing chromo or fluorescent mutant of a Cnidarian chromo- or  
CC fluorescent protein or its mutant. The protein is useful in applications  
CC involving nucleic acid encoding a chromo- or fluorescent protein and is  
CC useful for producing a chromo and/or fluorescent protein which involves  
CC growing the cell, whereby the protein is expressed, and isolating the  
CC protein substantially free of other proteins. The protein is useful in  
CC applications involving chromo- or fluorescent protein and is useful as  
CC PCR primers, hybridization probes, etc. The expression cassettes are  
CC useful for synthesizing related proteins. The chromoproteins are useful  
CC as coloring agents which are capable of imparting color or pigment to a  
CC particular composition of matter e.g. food compositions, pharmaceuticals,  
CC cosmetics, living organisms, e.g., animals and plants. The chromoproteins  
CC may also find use as labels in analyte detection assays, e.g. assays for  
CC biological analytes of interest and as selectable markers in recombinant  
CC DNA applications, e.g. the production of transgenic cells and organisms.  
CC The fluorescent proteins find use in a variety of different applications,  
CC e.g. in fluorescence resonance energy transfer (FRET) applications, as  
CC biosensors in prokaryotic and eukaryotic cells, in applications involving

CC the automated screening of arrays of cells expressing fluorescent  
CC reporting groups by using microscopic imaging and electronic analysis, as  
CC second messenger detectors, and in fluorescence activated cell sorting  
CC applications and as in vivo marker in animals. The fluorescent proteins  
CC also find use in protease cleavage assays. The proteins can also be used  
CC is assays to determine the phospholipid composition in biological  
CC membranes and as a fluorescent timer. The present sequence represents the  
CC wild-type DsRED.  
XX  
SQ Sequence 225 AA;

Alignment Scores: 3.01e-135 Length: 225  
Pred. No.: 1214.00 Matches: 225  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 99.92% Gaps: 0  
DB: 7

US-10-006-922A-11 (1-678) x ADN33979 (1-225)  
QY 1 ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTCCGATGGAAGGA 60  
Db 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
QY 61 ACGGTCAATGGGCACGAGTTTGAATAGAAAGCGGAAGGAGGGGAGGCCCATACGAAGGC 120  
Db 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGGACCTTTGCCATTTGCTGGGATATT 180  
Db 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
QY 181 TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCTGCCGACATACCA 240  
Db 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
QY 241 GACTATAAAAAGCTGTCTATTTTCCTGAAGGATTAAATGGGAAAGGTCATGAACCTTTGAA 300  
Db 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
QY 301 GACGGTGGCGTCGTTACTGTAAACCCAGGATTCCAGTTTGCAGGATGGCTGTTTCATCTAC 360  
Db 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
QY 361 AAGGTCAAGTTCATTGGCGTGAACCTTCCTCCGATGGACCTGTATGCAAAAGAAGACA 420  
Db 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGATCTCCTCGATGGCGTGTGAAAGGAGAG 480  
Db 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
QY 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT 540  
Db 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
QY 541 TACATGGCAAAGACCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAT 600  
Db 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
QY 601 ATAACAAGCCACACGAAGACTATACAATCGTTGAGCAGTATGAAAGACCGAGGGACGC 660  
Db 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
QY 661 CACCATCTGTTCCCTT 675  
Db 221 HisHisLeuPheLeu 225

RESULT 15  
ADI36421  
ID ADI36421 standard; protein; 225 AA.  
XX



AC ADI36421;  
XX DT 22-APR-2004 (first entry)  
XX Discosoma sp. red fluorescent protein (RED).  
DE Fluorobody; binding ligand; green fluorescent protein; GFP;  
XX target detection; red fluorescent protein; RED.  
KW Discosoma sp.  
XX

XX Key Location/Qualifiers  
FH Region 22..26  
FT /note= "Loop region"  
FT Region 37..39  
FT /note= "Loop region"  
FT Region 75..81  
FT /note= "Loop region"  
FT Region 100..103  
FT /note= "Loop region"  
FT Region 114..117  
FT /note= "Loop region"  
FT Region 153..156  
FT /note= "Loop region"  
FT Region 167..170  
FT /note= "Loop region"  
FT Region 185..192  
FT /note= "Loop region"  
FT Region 204..209  
FT /note= "Loop region"

XX US2003203355-A1.  
PN 30-OCT-2003.  
XX  
XX  
XX 24-APR-2002; 2002US-001322067.  
XX  
XX 24-APR-2002; 2002US-001322067.  
XX

PA (LALA-) LOS ALAMOS NAT LAB.  
PA (REGC ) UNIV CALIFORNIA.  
XX  
PI Bradbury AM, Zeytun A, Waldo GS;  
XX  
DR WPI; 2004-154325/15.  
DR N-PSDB; ADI36420.  
XX

PT Novel binding ligand with intrinsic fluorescence and comprising  
PT fluorescent protein having heterologous binding sites, useful for  
PT detecting target molecule.

XX Claim 2; SEQ ID NO 4; 23pp; English.  
XX  
CC The invention relates to binding ligands (fluorobodies) with intrinsic  
CC fluorescence, which comprises green fluorescent protein (GFP) having  
CC heterologous binding sites. The binding ligand is useful for detecting  
CC the target molecule and is efficiently detects the target molecule. The  
CC present sequence is Discosoma sp. red fluorescent protein (RED) used in  
CC the exemplification of the invention.  
XX

SQ Sequence 225 AA;

Alignment Scores:  
Pred. No.: 3.01e-135 Length: 225  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.92% Indels: 0  
DB: 8 Gaps: 0

US-10-006-922A-11 (1-678) x ADI36421 (1-225)

QY 1 ATGAGGTCTTCCAAGATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTCCGATGGAAGGA 60

Db 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
QY 61 ACGGTCATGGGCACGAGTTTGAATAGAACGCCAAGGAGAGGGGAGCCCATACGAAGGC 120  
Db 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
QY 121 CACAATACCGTAAAGCTTAAGGTAAACCAAGGGGGACCTTTGCCATTGCTTGGGATATT 180  
Db 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
QY 181 TTGTCAACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240  
Db 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
QY 241 GACTATAAAAAAGCTGTCTATTTCTCTGAAGGATTAAATGGGAAAGSGTCATGAACCTTTGAA 300  
Db 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
QY 301 GACGGTGGCGTCTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360  
Db 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
QY 361 AAGGTCAAAGTTCATTGGCGTGAACCTTCTCTCCGATGGACCTGTATGCAAAAAGAGACA 420  
Db 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGATCCTCGTGATGGCGTGTGAAAAGGAGAG 480  
Db 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
QY 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT 540  
Db 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
QY 541 TACATGGCAAGAAGCCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600  
Db 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrValAspSerLysLeuAsp 200  
QY 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGACGC 660  
Db 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
QY 661 CACCATCTGTTCCCTT 675  
Db 221 HisHisLeuPheLeu 225

Search completed: June 30, 2005, 15:58:40  
Job time : 140.5 secs



**THIS PAGE BLANK (u.s.r.e.)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: June 30, 2005, 15:54:05 ; Search time 137.5 Seconds  
(without alignments)  
3804.643 Million cell updates/sec

Title: US-10-006-922A-11  
Perfect score: 1215  
Sequence: 1 atgagggtcttccaagaatgt.....gccaccatctgttcctttaa 678

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1722976 seqs, 385795295 residues

Total number of hits satisfying chosen parameters: 3445952

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO spool\_p/US1006922/runat\_30062005\_110554\_9471/app query.fasta\_1.839  
-DB=Published Applications\_AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US1006922@cgn 1 1 199 @runat 30062005\_110554\_9471  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications\_AA:  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1214	99.9	225	9	US-09-999-745-67	Sequence 67, Appl
2	1214	99.9	225	10	US-09-866-538-12	Sequence 12, Appl
3	1214	99.9	225	10	US-09-794-308-12	Sequence 12, Appl
4	1214	99.9	225	10	US-09-865-291-12	Sequence 12, Appl
5	1214	99.9	225	13	US-10-006-922-12	Sequence 12, Appl
6	1214	99.9	225	14	US-10-081-864-8	Sequence 8, Appli
7	1214	99.9	225	14	US-10-121-258-1	Sequence 1, Appli
8	1214	99.9	225	14	US-10-315-920-2	Sequence 2, Appli
9	1214	99.9	225	15	US-10-132-067-4	Sequence 4, Appli
10	1214	99.9	225	15	US-10-370-570-56	Sequence 56, Appl
11	1214	99.9	225	15	US-10-406-618-32	Sequence 32, Appl
12	1214	99.9	225	16	US-10-433-640-13	Sequence 13, Appl
13	1214	99.9	225	16	US-10-311-030-7	Sequence 7, Appli
14	1214	99.9	225	16	US-10-845-484-3	Sequence 3, Appli
15	1214	99.9	225	16	US-10-885-988-12	Sequence 12, Appl
16	1214	99.9	225	17	US-10-656-029-22	Sequence 22, Appl
17	1214	99.9	225	17	US-10-857-622-12	Sequence 12, Appl
18	1214	99.9	225	17	US-10-505-486-26	Sequence 26, Appl
19	1214	99.9	487	15	US-10-343-977-1	Sequence 1, Appli
20	1214	99.9	487	17	US-10-343-977-1	Sequence 1, Appli
21	1214	99.9	506	15	US-10-343-977-2	Sequence 2, Appli
22	1214	99.9	506	17	US-10-343-977-2	Sequence 2, Appli
23	1214	99.9	547	15	US-10-343-977-3	Sequence 3, Appli
24	1214	99.9	547	17	US-10-343-977-3	Sequence 3, Appli
25	1210	99.6	226	16	US-10-311-030-9	Sequence 9, Appli
26	1210	99.6	240	14	US-10-152-296-2	Sequence 2, Appli
27	1210	99.6	240	16	US-10-739-656-2	Sequence 2, Appli
28	1210	99.6	545	14	US-10-214-932-52	Sequence 52, Appl
29	1210	99.6	548	14	US-10-214-932-76	Sequence 76, Appl
30	1207	99.3	225	14	US-10-121-258-20	Sequence 20, Appl
31	1207	99.3	225	14	US-10-315-920-4	Sequence 4, Appli
32	1196	98.4	236	16	US-10-314-936-2	Sequence 2, Appli
33	1196	98.4	236	16	US-10-314-936-4	Sequence 4, Appli
34	1196	98.4	236	19	US-11-021-014-2	Sequence 2, Appli
35	1196	98.4	236	19	US-11-021-014-4	Sequence 4, Appli
36	1195.5	98.4	226	16	US-10-724-178-12	Sequence 12, Appl
37	1191	98.0	225	14	US-10-315-920-6	Sequence 6, Appli
38	1186	97.6	225	15	US-10-442-148A-7	Sequence 7, Appli
39	1186	97.6	239	15	US-10-442-148A-8	Sequence 8, Appli
40	1160	95.5	225	14	US-10-121-258-4	Sequence 4, Appli
41	1153	94.9	225	14	US-10-121-258-24	Sequence 24, Appl
42	1132	93.2	225	15	US-10-423-688A-41	Sequence 41, Appl
43	1121	92.3	226	14	US-10-121-258-6	Sequence 6, Appli
44	1119	92.1	225	13	US-10-006-922-44	Sequence 44, Appl
45	1119	92.1	225	14	US-10-081-864-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1  
US-09-999-745-67  
; Sequence 67, Application US/09999745  
; Patent No. US20020157120A1  
; GENERAL INFORMATION:  
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: Tsien, Roger Y.  
; APPLICANT: Baird, Geoffrey  
; TITLE OF INVENTION: CIRCULARLY PERMUTED FLUORESCENT PROTEIN INDICATORS  
; FILE REFERENCE: REGEN1470-1  
; CURRENT APPLICATION NUMBER: US/09/999,745  
; CURRENT FILING DATE: 2001-10-23  
; PRIOR APPLICATION NUMBER: 09/316,920  
; PRIOR FILING DATE: 1999-05-21  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 67  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Discosoma sp.

```
US-09-999-745-67
Alignment Scores:
Pred. No.:      2.99e-128      Length:      225
Score:          1214.00      Matches:      225
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    99.92%      Indels:      0
DB:            9      Gaps:      0

US-10-006-922A-11 (1-678) x US-09-999-745-67 (1-225)
QY      1  ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGGA 60
      |||
Db      1  MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20

QY      61  ACGGTCAATGGGCACGAGTTTGAATAGAAAGCGCAAGGAGAGGGAGGCCATACGAAGGC 120
      |||
Db      21  ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40

QY      121  CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180
      |||
Db      41  HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60

QY      181  TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA 240
      |||
Db      61  LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80

QY      241  GACTATAAAAAGCTGTCATTTCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300
      |||
Db      81  AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100

QY      301  GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGATTTCAGGATGGCGTGTTCATCTAC 360
      |||
Db      101  AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120

QY      361  AAGGTCAAGTTTCATTTGGCGTGAACCTTTCTCCGATGGACCTGTTATGCAAAAAGAGACA 420
      |||
Db      121  LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140

QY      421  ATGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGTATGGCGTGTGAAAGGAGAG 480
      |||
Db      141  MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160

QY      481  ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTTCCTTAGTTGAATTCAAAAGTATT 540
      |||
Db      161  IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180

QY      541  TACATGGCAAAGAAGCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600
      |||
Db      181  TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200

QY      661  CACCATCTGTTCCCTT 675
      |||
Db      221  HisHisLeuPheLeu 225

RESULT 2
US-09-866-538-12
; Sequence 12, Application US/09866538
; Publication No. US20030032088A1
; GENERAL INFORMATION:
; APPLICANT:  TSIENT, Roger
; APPLICANT:  Campbell, Robert
; TITLE OF INVENTION:  NON-OLIGOMERIZING FLUORESCENT PROTEINS
; FILE REFERENCE:  REGEN1530-2
; CURRENT APPLICATION NUMBER:  US/09/866,538
; CURRENT FILING DATE:  2001-05-24
; NUMBER OF SEQ ID NOS: 29
```

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
US-09-866-538-12
Alignment Scores:
Pred. No.:      2.99e-128      Length:      225
Score:          1214.00      Matches:      225
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    99.92%      Indels:      0
DB:            10      Gaps:      0

US-10-006-922A-11 (1-678) x US-09-866-538-12 (1-225)
QY      1  ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGGA 60
      |||
Db      1  MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20

QY      61  ACGGTCAATGGGCACGAGTTTGAATAGAAAGCGCAAGGAGAGGGAGGCCATACGAAGGC 120
      |||
Db      21  ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40

QY      121  CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATT 180
      |||
Db      41  HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60

QY      181  TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA 240
      |||
Db      61  LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80

QY      241  GACTATAAAAAGCTGTCATTTCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300
      |||
Db      81  AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100

QY      301  GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGATTTCAGGATGGCGTGTTCATCTAC 360
      |||
Db      101  AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120

QY      361  AAGGTCAAGTTTCATTTGGCGTGAACCTTTCTCCGATGGACCTGTTATGCAAAAAGAGACA 420
      |||
Db      121  LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140

QY      421  ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGTATGGCGTGTGAAAGGAGAG 480
      |||
Db      141  MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160

QY      481  ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTTCATTCATTCATGTTGAATTCAAAAGTATT 540
      |||
Db      161  IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180

QY      541  TACATGGCAAAGAAGCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600
      |||
Db      181  TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200

QY      601  ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC 660
      |||
Db      201  IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220

QY      661  CACCATCTGTTCCCTT 675
      |||
Db      221  HisHisLeuPheLeu 225

RESULT 3
US-09-794-308-12
; Sequence 12, Application US/09794308
; Publication No. US20030170911A1
; GENERAL INFORMATION:
; APPLICANT:  REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT:  TSIENT, Roger
; APPLICANT:  ZACHARIAS, David
```

; APPLICANT: BAIRD, Geoffrey  
; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS  
; FILE REFERENCE: REGEN1530  
; CURRENT APPLICATION NUMBER: US/09/794,308  
; CURRENT FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 12  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Discosoma sp.  
US-09-794-308-12

Alignment Scores:  
Pred. No.: 2.99e-128 Length: 225  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.92% Indels: 0  
DB: 10 Gaps: 0

US-10-006-922A-11 (1-678) x US-09-794-308-12 (1-225)

QY 1 ATGAGGTCTTCCAAGATGTTATCAAGGAGTTCATGAGGTTTAAGGTTTCGCATGGAAGGA 60  
Db 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
QY 61 ACGGTCAATGGGCACGAGCTTTGAAATAGAACGCCGAGAGGGGAGGCCCATACGAAGGC 120  
Db 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTGGGATATT 180  
Db 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
QY 181 TTGTCACCAACAATTTCAAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA 240  
Db 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
QY 241 GACTATAAAAAGCTGTCATTTCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTGAA 300  
Db 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
QY 301 GACGGTGGCGTCTACTGTAACCCAGGATTTCCAGTTTGCAGGATGGCTGTTTCATCTAC 360  
Db 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
QY 361 AAGGTCAAGTTCAATTGGCGTGAACTTCTCCGATGGACCTGTTATGCAAAAGAGACA 420  
Db 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
Db 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
QY 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTTCATTTACCTAGTTGAATTCAAAAGTATT 540  
Db 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
QY 541 TACATGGCAAAGAAGCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600  
Db 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrValAspSerLysLeuAsp 200  
QY 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGACGC 660  
Db 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
QY 661 CACCATCTGTTCCCTT 675  
Db 221 HisHisLeuPheLeu 225

RESULT 4  
US-09-865-291-12

; Sequence 12, Application US/09865291  
; Publication No. US20030186229A1  
; GENERAL INFORMATION:  
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: TSUEN, Roger  
; APPLICANT: TING, Alice  
; APPLICANT: ZHANG, Jin  
; TITLE OF INVENTION: EMISSION RATIO-METRIC INDICATORS OF PHOSPHORYLATION  
; FILE REFERENCE: REGEN1550  
; CURRENT APPLICATION NUMBER: US/09/865,291  
; CURRENT FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 12  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Discosoma sp.  
US-09-865-291-12

Alignment Scores:  
Pred. No.: 2.99e-128 Length: 225  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.92% Indels: 0  
DB: 10 Gaps: 0

US-10-006-922A-11 (1-678) x US-09-865-291-12 (1-225)

QY 1 ATGAGGTCTTCCAAGATGTTATCAAGGAGTTCATGAGGTTTAAGGTTTCGCATGGAAGGA 60  
Db 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
QY 61 ACGGTCAATGGGCACGAGCTTTGAAATAGAACGCCGAGAGGGGAGGCCCATACGAAGGC 120  
Db 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGAGCCCTTTGCCATTTGCTGGGATATT 180  
Db 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
QY 181 TTGTCACCAACAATTTCAAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA 240  
Db 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
QY 241 GACTATAAAAAGCTGTCATTTCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTGAA 300  
Db 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
QY 301 GACGGTGGCGTCTACTGTAACCCAGGATTTCCAGTTTGCAGGATGGCTGTTTCATCTAC 360  
Db 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
QY 361 AAGGTCAAGTTCAATTGGCGTGAACTTCTCCGATGGACCTGTTATGCAAAAGAGACA 420  
Db 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
Db 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
QY 481 ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTTCATTTACCTAGTTGAATTCAAAAGTATT 540  
Db 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
QY 541 TACATGGCAAAGAAGCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT 600  
Db 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrValAspSerLysLeuAsp 200  
QY 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGACGC 660  
Db 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220



QY 661 CACCATCTGTTCCCTT 675  
Db 221 HisHisLeuPheLeu 225

RESULT 5  
US-10-006-922-12  
; Sequence 12, Application US/10006922  
; Publication No. US20020197676A1  
; GENERAL INFORMATION:  
; APPLICANT: Lukyanov, Sergey A  
; APPLICANT: Fradkov, Arcady F.  
; APPLICANT: Labas, Yulii A.  
; APPLICANT: Matz, Mikhail V.  
; APPLICANT: Tersikh, Alexey  
; TITLE OF INVENTION: No. US20020197676A1el Chromophores/Fluorophores and  
; TITLE OF INVENTION: Methods for Using the Same  
; FILE REFERENCE: CLON-035CIP  
; CURRENT APPLICATION NUMBER: US/10/006,922  
; CURRENT FILING DATE: 2001-12-04  
; PRIOR APPLICATION NUMBER: 09/120,330  
; PRIOR FILING DATE: 1998-12-11  
; PRIOR APPLICATION NUMBER: 09/457,898  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: 09/458,144  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: 09/458,477  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: 09/457,556  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: 09/444,338  
; PRIOR FILING DATE: 1999-11-19  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Discosoma species  
US-10-006-922-12

Alignment Scores:  
Pred. No.: 2,99e-128 Length: 225  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.92% Indels: 0  
DB: 13 Gaps: 0

US-10-006-922A-11 (1-678) x US-10-006-922-12 (1-225)

QY 1 ATGAGGTCTTCCAAGATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTCCGATGGAAGGA 60  
Db 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20

QY 61 ACGGTCAATGGGCACGAGTTTGAAATAGAGCGCAAGGAGGAGAGGGGCCATACGAAGGC 120  
Db 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40

QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTTGGCTTGGGATATT 180  
Db 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60

QY 181 TTGTCACCAACAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240  
Db 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80

QY 241 GACTATAAAAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAA 300  
Db 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100

QY 301 GACGGTGGCGTCGTTACTGTATAACCCAGGATTCCAGTTTGCCAGGATGGCTGTTTCATCTAC 360  
Db 101 AspGlyGlyValValThrValThrGlnAspSerSerSerLeuGlnAspGlyCysPheIleTyr 120

QY 361 AAGGTCAAGTTCATTGGCGTGAACTTTCTTCCGATGGACCTGTATGCAAAAAGAAGACA 420  
Db 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140

QY 421 ATGGGCTGGGAAGCCAGCACCTGAGCGTTTGATCCTCGTGATGGCGTGTGAAAGGAGAG 480  
Db 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160

QY 481 ATTCAATAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT 540  
Db 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180

QY 541 TACATGGCAAAAGACCTGTGCAGCTACCAGGTACTACTATGTGACTCCAAACTGGAT 600  
Db 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrValAspSerLysLeuAsp 200

QY 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCCGAGGACGC 660  
Db 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220

QY 661 CACCATCTGTTCCCTT 675  
Db 221 HisHisLeuPheLeu 225

RESULT 6  
US-10-081-864-8  
; Sequence 8, Application US/10081864  
; Publication No. US20030022287A1  
; GENERAL INFORMATION:  
; APPLICANT: Lukyanov, Sergey  
; APPLICANT: Lukyanov, Konstantin  
; APPLICANT: Yanushevich, Yuriy  
; APPLICANT: Savitsky, Alexandr  
; APPLICANT: Fradkov, Arcady  
; TITLE OF INVENTION: No. US20030022287A1 Aggregating Fluorescent Proteins and  
; TITLE OF INVENTION: Methods for Using the Same  
; FILE REFERENCE: CLON-067  
; CURRENT APPLICATION NUMBER: US/10/081,864  
; CURRENT FILING DATE: 2002-06-19  
; PRIOR APPLICATION NUMBER: 10/006,922  
; PRIOR FILING DATE: 2001-12-04  
; PRIOR APPLICATION NUMBER: 60/270,983  
; PRIOR FILING DATE: 2001-02-21  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Discosoma sp.  
US-10-081-864-8

Alignment Scores:  
Pred. No.: 2,99e-128 Length: 225  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.92% Indels: 0  
DB: 14 Gaps: 0

US-10-006-922A-11 (1-678) x US-10-081-864-8 (1-225)

QY 1 ATGAGGTCTTCCAAGAAATGTTATCAAGGAGTTTCATGAGGTTTAAGTTCGCATGGAAGGA 60  
Db 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20

QY 61 ACGGTCAATGGGCACGAGTTTGAAATAGAGCGCAAGGAGGAGAGGGGCCATACGAAGGC 120  
Db 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40

QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTTGGCTTGGGATATT 180  
Db 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60

QY	181	TTGTCAACACAATTTCAGTATGGAAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA	240
Db	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
QY	241	GACTATAAAAAGCTGTTCATTTCCCTGAAGGATTAAATGGGAAAGGTCATGAACTTTGAA	300
Db	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
QY	301	GACGGTGGCGTCGTTACTGTAAACCCAGGATTCACAGTTTGCAGGATGGCTGTTTCATCTAC	360
Db	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
QY	361	AAGGTCAAAGTTTCATTGGCGTGAACCTTCTCTCCGATGGACCTGTTATGCAAAAAGAAGACA	420
Db	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
QY	421	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGGTGTTGAAAGGAGAG	480
Db	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
QY	481	ATTCAATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT	540
Db	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
QY	541	TACATGGCAAAGACCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAT	600
Db	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrValAspSerLysLeuAsp	200
QY	601	ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCAGGGGACGC	660
Db	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
QY	661	CACCATCTGTTCCCTT	675
Db	221	HisHisLeupheLeu	225

## RESULT 7

```

US-10-121-258-1
; Sequence 1, Application US/10121258
; Publication No. US20030059835A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CPI
; CURRENT APPLICATION NUMBER: US/10/121,258
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(225)
; OTHER INFORMATION: wild-type DsRed
US-10-121-258-1

```

Alignment Scores:			
Pred. No.:	2,99e-128	Length:	225
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	99.92%	Indels:	0
DB:	14	Gaps:	0

US-10-006-922A-11 (1-678) x US-10-121-258-1 (1-225)

QY	1	ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTCATGAGGTTTAAAGGTTTCGCATGGAAAGGA	60
Db	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
QY	61	ACGGTCAATGGGCAAGAGTTTGAATAGAAAGCGCAAGAGAGGGGAGGCCATACGAAGGC	120
Db	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly	40
QY	121	CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGGACCTTTGCCATTTGCTTGGGATATT	180
Db	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle	60
QY	181	TTGTCAACCAATTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA	240
Db	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
QY	241	GACTATAAAAGCTGTTCATTTCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAA	300
Db	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
QY	301	GACGGTGGCGTCGTTACTGTATAACCCAGGATTCCAGTTTGCAGGATGGCTGTTTCATCTAC	360
Db	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
QY	361	AAGGTCAAGTTCAATTGGCGTGAACTTTCCTTCCGATGGACCTGTTATGCAAAAGAAGACA	420
Db	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
QY	421	ATGGGCTGGGAAGCCAGCACCTGAGCGTTTGTATCCTCGTGATGGCGTGTGTAAGGAGAG	480
Db	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
QY	481	ATTCAATAAGGCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT	540
Db	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
QY	541	TACATGGCAAAGRAGCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGAT	600
Db	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrValAspSerLysLeuAsp	200
QY	601	ATAACAAGCCACACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC	660
Db	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
QY	661	CACCATCTGTTCCCTT	675
Db	221	HisHisLeuPheLeu	225

## RESULT 8

```

US-10-315-920-2
; Sequence 2, Application US/10315920
; Publication No. US20030175809A1
; GENERAL INFORMATION:
; APPLICANT: Fradkov, Arcady Fedorovich
; APPLICANT: Terskikh, Alexey
; TITLE OF INVENTION: FLUORESCENT TIMER PROTEINS AND METHODS
; TITLE OF INVENTION: FOR THEIR USE
; FILE REFERENCE: CLON-077CIP
; CURRENT APPLICATION NUMBER: US/10/315,920
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/211,607
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: PCT/US01/19097
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
US-10-315-920-2

```





; TITLE OF INVENTION: FLUOROPHORE COMPLEMENTATION PRODUCTS
; FILE REFERENCE: 3759-0130P
; CURRENT APPLICATION NUMBER: US/10/370,570
; CURRENT FILING DATE: 2003-02-24
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 56
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
US-10-370-570-56

Alignment Scores:
Pred. No.: 2.99e-128 Length: 225
Score: 1214.00 Matches: 225
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.92% Indels: 0
DB: 15 Gaps: 0

US-10-006-922A-11 (1-678) x US-10-370-570-56 (1-225)

QY 1 ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGGA 60
Db 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20

QY 61 ACGGTCAATGGGCACGAGTTTGAATAGAAAGCGAAGGAGAGGGAGGCCATACGAAGGC 120
Db 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40

QY 121 CACAATACCGTAAAGCTTAAAGGTAACCAAGGGGGACCTTTGCCATTTGCTGGGATATT 180
Db 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60

QY 181 TTGTCACCACAAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA 240
Db 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80

QY 241 GACTATAAAAAGCTGTTCATTTCTGAAGGATTTAAATGGAAAGGGTCATGAACCTTTGAA 300
Db 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100

QY 301 GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360
Db 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120

QY 361 AAGGTCAAGTTTCATTTGGCGTGAACTTTCCTTCGATGGACCTGTTTGAAGGAGAG 420
Db 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140

QY 421 ATTCATAAGGCTCTGAAGCTGAAAGACCGGTGCTACCTAGTTGAATTCAAAAGTATT 540
Db 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180

QY 541 TACATGGCAAAGACCTGTGCAGCTACCGAGGTACTACTATGTTGACTCCAAACTGGAT 600
Db 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrValAspSerLysLeuAsp 200

QY 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGACGAGTATGAAAGAACCGAGGGACGC 660
Db 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220

QY 661 CACCATCTGTTCCCTT 675
Db 221 HisHisLeuPheLeu 225

RESULT 11
US-10-406-618-32
; Sequence 32, Application US/10406618

; Publication No. US20030219814A1
; GENERAL INFORMATION:
; APPLICANT: Wan, David Chi-Cheong
; APPLICANT: Ip, Denis Tsz-Ming
; APPLICANT: The Chinese University of Hong Kong
; TITLE OF INVENTION: No. US20030219814A1el Fluorescent Proteins
; FILE REFERENCE: 016285-34-1
; CURRENT APPLICATION NUMBER: US/10/406,618
; CURRENT FILING DATE: 2003-04-02
; PRIOR APPLICATION NUMBER: US 60/387,968
; PRIOR FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: US 60/370,598
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Spontaneously
; OTHER INFORMATION: fluorescent protein DsRed.
US-10-406-618-32

Alignment Scores:
Pred. No.: 2.99e-128 Length: 225
Score: 1214.00 Matches: 225
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.92% Indels: 0
DB: 15 Gaps: 0

US-10-006-922A-11 (1-678) x US-10-406-618-32 (1-225)

QY 1 ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGGA 60
Db 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20

QY 61 ACGGTCAATGGGCACGAGTTTGAATAGAAAGCGAAGGAGAGGGAGGCCATACGAAGGC 120
Db 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40

QY 121 CACAATACCGTAAAGCTTAAAGGTAACCAAGGGGGACCTTTGCCATTTGCTGGGATATT 180
Db 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60

QY 181 TTGTCACCACAAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA 240
Db 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80

QY 241 GACTATAAAAAGCTGTTCATTTCTGAAGGATTTAAATGGAAAGGGTCATGAACCTTTGAA 300
Db 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100

QY 301 GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360
Db 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120

QY 361 AAGGTCAAGTTTCATTTGGCGTGAACTTTCCTTCGATGGACCTGTTTGAAGGAGAGACA 420
Db 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140

QY 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTTCGATGGCGTGTGTAAGGAGAG 480
Db 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160

QY 481 ATTCATAAGGCTCTGAAGCTGAAAGACCGGTGGTTCATTACCTAGTTGAATTCAAAAGTATT 540
Db 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180

QY 541 TACATGGCAAAGACCTGTGCAGCTACCGAGGTACTACTATGTTGACTCCAAACTGGAT 600
Db 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrValAspSerLysLeuAsp 200





QY 301 GACGGTGGCGTGTACTGTAAACCCAGGATTCCAGTTTGCAGGATGGCTGTTTCATCTAC 360  
| | | | |  
Db 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
| | | | |  
QY 361 AAGGTCAAGTTCATTGGCGTGAACTTTCCTTCGGATGGACCTGTTATGCAAAAAGAGACA 420  
| | | | |  
Db 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
| | | | |  
QY 421 ATGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAAGGAGAG 480  
| | | | |  
Db 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
| | | | |  
QY 481 ATTCATAAGGCTCTGAAGCTGAAAGACCGTGGTCATTACCTAGTTGAATTCAAAAGTATT 540  
| | | | |  
Db 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
| | | | |  
QY 541 TACATGGCAAAAGAGCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAT 600  
| | | | |  
Db 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
| | | | |  
QY 601 ATAACAAGCCACAAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGGAGCGC 660  
| | | | |  
Db 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
| | | | |  
QY 661 CACCATCTGTTCCCTT 675  
| | | | |  
Db 221 HisHisLeuPheLeu 225  
| | | | |

RESULT 14

US-10-845-484-3  
; Sequence 3, Application US/10845484  
; Publication No. US20040248180A1  
; GENERAL INFORMATION:  
; APPLICANT: Bulina, Maria  
; APPLICANT: Chudakov, Dmitry  
; APPLICANT: Lukyanov, Konstantin  
; TITLE OF INVENTION: Mutant Chromaphores/Flourophores and  
; TITLE OF INVENTION: Methods for Making and Using the Same  
; FILE REFERENCE: CLON 092  
; CURRENT APPLICATION NUMBER: US/10/845,484  
; CURRENT FILING DATE: 2004-05-12  
; PRIOR APPLICATION NUMBER: 60/343128  
; PRIOR FILING DATE: 2001-12-26  
; PRIOR APPLICATION NUMBER: PCT/US02/41418  
; PRIOR FILING DATE: 2002-12-23  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Discosoma sp.  
US-10-845-484-3

Alignment Scores:  
Pred. No.: 2.99e-128 Length: 225  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.92% Indels: 0  
DB: 16 Gaps: 0

US-10-006-922A-11 (1-678) x US-10-845-484-3 (1-225)

QY 1 ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTTCGCATGGAAGGA 60  
| | | | |  
Db 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
| | | | |  
QY 61 ACGGTCAATGGCACGAGTTTGAATAGAACGCAAGGAGAGGGGAGGCCATACGAAGGC 120  
| | | | |  
Db 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
| | | | |  
QY 121 CACAATACCGTAAAGCTTAAGGTAAACCAAGGGGGACCTTTGCCATTTGCTGGGATATT 180  
| | | | |

Db 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
| | | | |  
QY 181 TTGTCACCACAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTCCGACATACCA 240  
| | | | |  
Db 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
| | | | |  
QY 241 GACTATAAAAAGCTGTTCATTTCTCTGAAGGATTAAATGGGAAAGGGTTCATGAACCTTTGAA 300  
| | | | |  
Db 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
| | | | |  
QY 301 GACGGTGGCGTGTACTGTAAACCCAGGATTCAGTTCAGGATGGCTGTTTCATCTAC 360  
| | | | |  
Db 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
| | | | |  
QY 361 AAGGTCAAGTTCATTGGCGTGAACCTTTCCTTCCGATGGACCTGTTATGCAAAAAGAGACA 420  
| | | | |  
Db 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
| | | | |  
QY 421 ATGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAAGGAGAG 480  
| | | | |  
Db 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
| | | | |  
QY 481 ATTCATAAGGCTCTGAAGCTGAAAGACCGTGGTCATTACCTAGTTGAATTCAAAAGTATT 540  
| | | | |  
Db 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
| | | | |  
QY 541 TACATGGCAAAAGACCTGTGCAGCTACCAAGGTTACTACTATGTTGATCCCAAACTGGAT 600  
| | | | |  
Db 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
| | | | |

RESULT 15

US-10-885-988-12  
; Sequence 12, Application US/10885988  
; Publication No. US20040259165A1  
; GENERAL INFORMATION:  
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: TSJEN, Roger  
; APPLICANT: Campbell, Robert  
; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS  
; FILE REFERENCE: REGEN1530-2  
; CURRENT APPLICATION NUMBER: US/10/885,988  
; CURRENT FILING DATE: 2004-07-06  
; PRIOR APPLICATION NUMBER: US/09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 12  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Discosoma sp.  
US-10-885-988-12

Alignment Scores:  
Pred. No.: 2.99e-128 Length: 225  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.92% Indels: 0  
DB: 16 Gaps: 0

US-10-006-922A-11 (1-678) x US-10-885-988-12 (1-225)

QY 1 ATGAGGTCTTCCAAGAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTTCGCATGGAAGGA 60  
| | | | |  
Db 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
| | | | |

QY	61	ACGGTCAATGGGCACGAGTTTGAAATAGAACGCCGAGAGAGGGGAGGCCCATACGAAGGC	120
Db	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
QY	121	CACAATACCGTAAAGCTTAAGGTAAACCAAGGGGGACCTTTGCCATTTTGGTGATATT	180
Db	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle	60
QY	181	TTGTCACCACAAATTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA	240
Db	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
QY	241	GACTATAAAAAAGCTGTCATTTCTGAAGGATTTAAATGGGAAAGGGTCATGAACTTTGAA	300
Db	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
QY	301	GACGGTGGCGTCGTACTGTAAACCCAGGATTCAGTTTGACAGGATGGCTGTTTCATCTAC	360
Db	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
QY	361	AAGGTCAAGTTTCATTGGCGTGAACTTTCCTTCCGATGGACCTGTTATGCAAAAAGAAGACA	420
Db	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
QY	421	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAG	480
Db	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
QY	481	ATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTGCTACCTAGTTGAATTCAAAAAGTATT	540
Db	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
QY	541	TACATGGCAAAGCCTGTGCAGCTACCAGGCTACTACTATGTTGACTCCAAAACCTGGAT	600
Db	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrValAspSerLysLeuAsp	200
QY	601	ATAACAAGCCACACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC	660
Db	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
QY	661	CACCATCTGTTCCTTT	675
Db	221	HisHisLeuPheLeu	225

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: June 30, 2005, 15:51:15 ; Search time 32 Seconds  
(without alignments)  
3163.254 Million cell updates/sec

Title: US-10-006-922A-11  
Perfect score: 1215  
Sequence: 1 atgagggtcttccaagaatgt.....gccaccatctgttcctttaa 678

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO spool\_p/US10006922/runat 30062005 110553 9450/app query.fasta\_1.839  
-DB=Issued\_Patents\_AA -QFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10006922@cgn\_1\_1\_46@runat 30062005 110553 9450 -NCPUS=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PCITUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1210	99.6	240	4	US-10-152-296-2
2	485.5	40.0	238	3	US-09-277-716-16
3	485.5	40.0	238	4	US-09-609-161B-16
4	485.5	40.0	238	4	US-09-626-581D-65
5	485.5	40.0	238	4	US-09-415-765B-65
6	485.5	40.0	238	4	US-09-626-580C-65
7	473.5	39.0	224	4	US-09-977-897-3
8	473.5	39.0	224	4	US-09-977-897-13
9	473.5	39.0	225	4	US-09-977-897-14
10	473.5	39.0	226	4	US-09-977-897-15
11	473.5	39.0	227	4	US-09-977-897-16
12	473.5	39.0	228	4	US-09-977-897-17
					Sequence 2, Appli
					Sequence 16, Appl
					Sequence 16, Appl
					Sequence 65, Appl
					Sequence 65, Appl
					Sequence 3, Appli
					Sequence 13, Appl
					Sequence 14, Appl
					Sequence 15, Appl
					Sequence 16, Appl
					Sequence 17, Appl

13	473.5	39.0	229	4	US-09-977-897-18	Sequence 18, Appl
14	473.5	39.0	230	4	US-09-977-897-19	Sequence 19, Appl
15	473.5	39.0	231	4	US-09-977-897-20	Sequence 20, Appl
16	473.5	39.0	232	4	US-09-977-897-21	Sequence 21, Appl
17	473.5	39.0	233	4	US-09-977-897-22	Sequence 22, Appl
18	473.5	39.0	234	4	US-09-977-897-9	Sequence 9, Appli
19	473.5	39.0	234	4	US-09-977-897-23	Sequence 23, Appl
20	473.5	39.0	235	4	US-09-977-897-8	Sequence 8, Appli
21	473.5	39.0	235	4	US-09-977-897-24	Sequence 24, Appl
22	473.5	39.0	236	4	US-09-977-897-7	Sequence 7, Appli
23	473.5	39.0	236	4	US-09-977-897-25	Sequence 25, Appl
24	473.5	39.0	237	4	US-09-977-897-6	Sequence 6, Appli
25	473.5	39.0	237	4	US-09-977-897-26	Sequence 26, Appl
26	473.5	39.0	238	3	US-09-277-716-32	Sequence 32, Appl
27	473.5	39.0	238	4	US-09-609-161B-32	Sequence 32, Appl
28	473.5	39.0	238	4	US-09-839-650-3	Sequence 3, Appli
29	473.5	39.0	238	4	US-09-977-897-5	Sequence 5, Appli
30	473.5	39.0	238	4	US-09-977-897-27	Sequence 27, Appl
31	473.5	39.0	239	4	US-09-977-897-2	Sequence 2, Appli
32	473	38.9	219	4	US-09-977-897-4	Sequence 4, Appli
33	473	38.9	231	4	US-09-977-897-12	Sequence 12, Appl
34	473	38.9	232	4	US-09-977-897-11	Sequence 11, Appl
35	473	38.9	233	4	US-09-977-897-10	Sequence 10, Appl
36	217.5	17.9	238	4	US-09-023-946B-23	Sequence 23, Appl
37	216.5	17.8	238	4	US-09-023-946B-28	Sequence 28, Appl
38	214.5	17.7	238	2	US-08-818-604-32	Sequence 32, Appl
39	214.5	17.7	238	3	US-08-819-612-22	Sequence 22, Appl
40	214.5	17.7	238	4	US-09-316-919-3	Sequence 3, Appli
41	214.5	17.7	238	4	US-09-346-946-32	Sequence 32, Appl
42	214.5	17.7	238	4	US-09-023-946B-22	Sequence 22, Appl
43	214.5	17.7	238	4	US-09-316-920A-3	Sequence 3, Appli
44	214.5	17.7	238	4	US-09-872-364-22	Sequence 22, Appl
45	213.5	17.6	240	4	US-09-129-192C-49	Sequence 49, Appl

ALIGNMENTS

RESULT 1  
US-10-152-296-2  
; Sequence 2, Application US/10152296  
; Patent No. 6723537  
; GENERAL INFORMATION:  
; APPLICANT: Peelle, Beau  
; APPLICANT: Rigel Pharmaceuticals, Incorporated  
; TITLE OF INVENTION: Directed Evolution of Protein in Mammalian Cells  
; FILE REFERENCE: 021044-000110US  
; CURRENT APPLICATION NUMBER: US/10/152,296  
; CURRENT FILING DATE: 2002-12-10  
; PRIOR APPLICATION NUMBER: US 60/291,871  
; PRIOR FILING DATE: 2001-05-18  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 240  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:mammalian  
; OTHER INFORMATION: codon-optimized variant (DSRED) of Discosoma sp.  
; OTHER INFORMATION: "red" red fluorescent protein (RFP)  
US-10-152-296-2

Alignment Scores:  
Pred. No.: 1.16e-137 Length: 240  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.59% Indels: 0  
DB: 4 Gaps: 0

US-10-006-922A-11 (1-678) x US-10-152-296-2 (1-240)

QY 1 ATGAGGTCCTCCAAGAATGTTATCAAGGAGTTCATGAGGTTTAAGTTCCGATGGAAGGA 60



Db 2 ValArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 21  
QY 61 ACGGTCAATGGGCACGAGTTTGAAATAGAAAGCGAAGAGAGAGGGGAGGCATACGAAGGC 120  
Db 22 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 41  
QY 121 CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTGGGATATT 180  
Db 42 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspile 61  
QY 181 TTGTCACCCACAATTTCAAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA 240  
Db 62 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspilePro 81  
QY 241 GACTATAAAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAGGTCATGAACTTTGAA 300  
Db 82 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 101  
QY 301 GACGGTGGCGTCGTTACTGTAAACCCAGGATTCAGTTTCAGGATGGCGTGTTCATCTAC 360  
Db 102 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 121  
QY 361 AAGGTCAAGTTCAATGGCGTGAACTTTCCTCCGATGGACCTGTTATGCAAAAGAGACA 420  
Db 122 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 141  
QY 421 ATGGGCTGGGAAGCCAGCACACTGAGCGGTTTGATTCCTCGTGATGGCGTGTGAAAGGAGAG 480  
Db 142 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 161  
QY 481 ATTCAATAGGCTCTGAAGCTGAAAGACGCGTGCTCATTTACCTAGTTGAATTCAAAAGTATT 540  
Db 162 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 181  
QY 541 TACATGGCAAAGAGCCTGTGCAGCTACCAGGCTACTACTATGTTGACTCCAAACTGGAT 600  
Db 182 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 201  
QY 601 ATAACAAGCCACACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGGACGC 660  
Db 202 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 221  
QY 661 CACCATCTGTTTCCTT 675  
Db 222 HisHisLeuPheLeu 226

RESULT 2  
US-09-277-716-16  
; Sequence 16, Application US/09277716A  
; Patent No. 6232107  
; GENERAL INFORMATION:  
; APPLICANT: Bryan, Bruce  
; APPLICANT: Szent-Gyorgyi, Christopher  
; APPLICANT: PROLUME, LTD.  
; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE  
; CURRENT APPLICATION NUMBER: US/09/277,716A  
; EARLIER FILING DATE: 1999-03-26  
; EARLIER APPLICATION NUMBER: 60/102,939  
; EARLIER FILING DATE: 1998-10-01  
; EARLIER APPLICATION NUMBER: 60/089,367  
; EARLIER FILING DATE: 1998-06-15  
; EARLIER APPLICATION NUMBER: 60/079,624  
; EARLIER FILING DATE: 1998-03-27  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 16  
; LENGTH: 238  
; TYPE: PRT  
; ORGANISM: Renilla mulleri  
; FEATURE:  
; OTHER INFORMATION: Renilla mulleri Green Fluorescent Protein (GFP)

Alignment Scores:  
Pred. No.: 7.56e-50 Length: 238  
Score: 485.50 Matches: 97  
Percent Similarity: 64.22% Conservative: 43  
Best local Similarity: 44.50% Mismatches: 71  
Query Match: 39.96% Indels: 7  
DB: 3 Gaps: 3

US-10-006-922A-11 (1-678) x US-09-277-716-16 (1-238)

QY 10 TCCAAGAATGTTATCAAG-----GAGTTTCATGAGGTTTAAGGTTTCGCATG 54  
Db 2 SerLysGlnIleLeuLysAsnThrCysLeuGlnGluValMetSerTyrLysValAsnLeu 21  
QY 55 GAAGGAACGGTCAATGGGCACGAGTTTGAAATAGAAAGCGAAGAGAGGGAGGCCATAC 114  
Db 22 GluGlyIleValAsnAsnHisValPheThrMetGluGlyCysGlyLysGlyAsnIleLeu 41  
QY 115 GAAGGCCACAATACCGTAAAGCTTAAGGTAACCAAGGGGGGACCTTTGCCATTTGCTTGG 174  
Db 42 PheGlyAsnGlnLeuValGlnIleArgValThrLysGlyAlaProLeuProPheAlaPhe 61  
QY 175 GATATTTGTCACCACAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGAC 234  
Db 62 AspIleValSerProAlaPheGlnTyrGlyAsnArgThrPheThrLysTyrProAsnAsp 81  
QY 235 ATACCAGACTATAAAAAGCTGTCATTTCTGAAGGATTTAAATGGGAAAGGGTCATGAAC 294  
Db 82 IleSerAspTyrPheIleGlnSerPheProAlaGlyPheMetTyrGluArgThrLeuArg 101  
QY 295 TTTGAAGACGGTGGCGTCGTTACTGTAACCCAGGATTCAGTTTGACAGGATGGCTGTTTC 354  
Db 102 TyrGluAspGlyGlyLeuValGluIleArgSerAspIleAsnLeuIleGluAspLysPhe 121  
QY 355 ATCTACAAGGTCAAGTTCATTGGCGTGAACTTTTCCTTCCGATGGACCTGTTATGCAAAAG 414  
Db 122 ValTyrArgValGluTyrLysGlySerAsnPheProAspAspGlyProValMetGlnLys 141  
QY 415 AAGACAATGGGCTGGGAAGCCAGCACTGAGCGTTTGATATCCTTCGTGATGGCGTGTGAAA 474  
Db 142 ThrIleLeuGlyIleGluProSerPheGluAlaMetTyrMetAsnAsnGlyValLeuVal 161  
QY 475 GGAGAGATTCATAAGGCTCTGAAGCTGAAAGACGGTGGTCACTACCTAGTTGAATTCAAA 534  
Db 162 GlyGluValIleLeuValTyrLysLeuAsnSerGlyLysTyrTyrSerCysHisMetLys 181  
QY 535 AGTATTTACATGGCAAAGAAGCCCTGTG---CAGTACCAGGGTACTACTATGTTGACTCC 591  
Db 182 ThrLeuMetLysSerLysGlyValValLysGluPheProSerTyrHisPheIleGlnHis 201  
QY 592 AAACCTGGATATAACAAGCCACACGAAGACTATACAATCGTTTGACAGCATATGAA 645  
Db 202 ArgLeuGluLysThr--TyrValGluAspGlyGlyPheValGluGlnHisGlu 218

RESULT 3  
US-09-609-161B-16  
; Sequence 16, Application US/09609161B  
; Patent No. 6436682  
; GENERAL INFORMATION:  
; APPLICANT: Bryan, Bruce  
; APPLICANT: Szent-Gyorgyi, Christopher  
; APPLICANT: PROLUME, LTD.  
; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE LUC  
; TITLE OF INVENTION: AND FLUORESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIG  
; TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS  
; FILE REFERENCE: 24729-121B  
; CURRENT APPLICATION NUMBER: US/09/609,161B  
; CURRENT FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: 09/277,716  
; PRIOR FILING DATE: 1999-03-26  
; PRIOR APPLICATION NUMBER: 60/102,939  
; PRIOR FILING DATE: 1998-10-01

; PRIOR APPLICATION NUMBER: 60/089,367
; PRIOR FILING DATE: 1998-06-15
; PRIOR APPLICATION NUMBER: 60/079,624
; PRIOR FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Renilla mulleri
; FEATURE:
; OTHER INFORMATION: Renilla mulleri Green Fluorescent Protein (GFP)
US-09-609-161B-16

Alignment Scores:
Pred. No.: 7.56e-50 Length: 238
Score: 485.50 Matches: 97
Percent Similarity: 64.22% Conservative: 43
Best Local Similarity: 44.50% Mismatches: 71
Query Match: 39.96% Indels: 7
DB: 4 Gaps: 3

US-10-006-922A-11 (1-678) x US-09-609-161B-16 (1-238)

QY 10 TCCAAGAATGTTATCAAG-----GAGTTCATGAGGTTTAAGGTTGCGCATG 54
Db 2 SerLysGlnIleLeuLysAsnThrCysLeuGlnGluValMetSerTyrLysValAsnLeu 21

QY 55 GAAGGAACGGTCAATGGGCACGAGTTTGAATAGAGCGAAGGAGAGGGGAGGCCATAC 114
Db 22 GluGlyIleValAsnAsnHisValPheThrMetGluGlyCysGlyLysGlyAsnIleLeu 41

QY 115 GAAGGCCACAATACCGTTAAAGCTTAAGGTAAACCAAGGGGGACCTTTGCCATTGCTTGG 174
Db 42 PheGlyAsnGlnLeuValGlnIleArgValThrLysGlyAlaProLeuProPheAlaPhe 61

QY 175 GATATTTGTCCACCAATTTCAAGTATGGAACCAAGGTATATGTCAAGCACCCCTGCCGAC 234
Db 62 AspIleValSerProAlaPheGlnTyrGlyAsnArgThrPheThrLysTyrProAsnAsp 81

QY 235 ATACCAGACTATAAAAGCTGTCTATTCCTGAAGGATTTAAATGGGAAGGTCATGAAC 294
Db 82 IleSerAspTyrPheIleGlnSerPheProAlaGlyPheMetTyrGluArgThrLeuArg 101

QY 295 TTTGAAGACGGTGGCGTCTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTC 354
Db 102 TyrGluAspGlyGlyLeuValGluIleArgSerAspIleAsnLeuIleGluAspLysPhe 121

QY 355 ATCTACAAGTCAAGTTCATTGGCGTGAACCTTCCGATGGACCTGTTATGCAAAAG 414
Db 142 ThrIleLeuGlyIleGluProSerPheGluAlaMetTyrMetAsnAsnGlyValLeuVal 161

QY 475 GGAGAGATTTCATAGGCTCTGAAGCTGAAAGACGGTGGTCAATTACCTAGTTGAATTCAAA 534
Db 162 GlyGluValIleLeuValTyrLysLeuAsnSerGlyLysTyrTyrSerCysHisMetLys 181

QY 535 AGTATTTACATGGCAAAGAGCCCTGTG---CAGTACCAGGGTACTACTATGTTGACTCC 591
Db 182 ThrLeuMetLysSerLysGlyValValLysGluPheProSerTyrHisPheIleGlnHis 201

QY 592 AAACCTGGATATAACAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAA 645
Db 202 ArgLeuGluLysThr---TyrValGluAspGlyGlyPheValGluGlnHisGlu 218

RESULT 4
US-09-626-581D-65
; Sequence 65, Application US/09626581D
; Patent No. 6548249
; GENERAL INFORMATION:

; APPLICANT: Anderson, David
; TITLE OF INVENTION: Fusions of Scaffold Proteins with Random Peptide
; FILE REFERENCE: A-66900-3/RMS
; CURRENT APPLICATION NUMBER: US/09/626,581D
; CURRENT FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 09/169,015
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 09/415,765
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 65
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Renilla muelleri
US-09-626-581D-65

Alignment Scores:
Pred. No.: 7.56e-50 Length: 238
Score: 485.50 Matches: 97
Percent Similarity: 64.22% Conservative: 43
Best Local Similarity: 44.50% Mismatches: 71
Query Match: 39.96% Indels: 7
DB: 4 Gaps: 3

US-10-006-922A-11 (1-678) x US-09-626-581D-65 (1-238)

QY 10 TCCAAGAATGTTATCAAG-----GAGTTCATGAGGTTTAAGGTTGCGCATG 54
Db 2 SerLysGlnIleLeuLysAsnThrCysLeuGlnGluValMetSerTyrLysValAsnLeu 21

QY 55 GAAGGAACGGTCAATGGGCACGAGTTTGAATAGAGCGAAGGAGAGGGGAGGCCATAC 114
Db 22 GluGlyIleValAsnAsnHisValPheThrMetGluGlyCysGlyLysGlyAsnIleLeu 41

QY 115 GAAGGCCACAATACCGTTAAAGCTTAAGGTAAACCAAGGGGGACCTTTGCCATTGCTTGG 174
Db 42 PheGlyAsnGlnLeuValGlnIleArgValThrLysGlyAlaProLeuProPheAlaPhe 61

QY 175 GATATTTGTCCACCAATTTCAAGTATGGAACCAAGGTATATGTCAAGCACCCCTGCCGAC 234
Db 62 AspIleValSerProAlaPheGlnTyrGlyAsnArgThrPheThrLysTyrProAsnAsp 81

QY 235 ATACCAGACTATAAAAGCTGTCTATTCCTGAAGGATTTAAATGGGAAGGTCATGAAC 294
Db 82 IleSerAspTyrPheIleGlnSerPheProAlaGlyPheMetTyrGluArgThrLeuArg 101

QY 295 TTTGAAGACGGTGGCGTCTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTC 354
Db 102 TyrGluAspGlyGlyLeuValGluIleArgSerAspIleAsnLeuIleGluAspLysPhe 121

QY 355 ATCTACAAGTCAAGTTCATTGGCGTGAACCTTCCGATGGACCTGTTATGCAAAAG 414
Db 122 ValTyrArgValGluTyrLysGlySerAsnPheProAspAspGlyProValMetGlnLys 141

QY 415 AAGACAATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGTGCGGTTGAAA 474
Db 142 ThrIleLeuGlyIleGluProSerPheGluAlaMetTyrMetAsnAsnGlyValLeuVal 161

QY 475 GGAGAGATTTCATAGGCTCTGAAGCTGAAAGACGGTGGTCAATTACCTAGTTGAATTCAAA 534
Db 162 GlyGluValIleLeuValTyrLysLeuAsnSerGlyLysTyrTyrSerCysHisMetLys 181

QY 535 AGTATTTACATGGCAAAGAGCCCTGTG---CAGTACCAGGGTACTACTATGTTGACTCC 591
Db 182 ThrLeuMetLysSerLysGlyValValLysGluPheProSerTyrHisPheIleGlnHis 201

QY 592 AAACCTGGATATAACAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAA 645
Db 202 ArgLeuGluLysThr---TyrValGluAspGlyGlyPheValGluGlnHisGlu 218

RESULT 5





Db 182 ThrLeuMetLysSerLysGlyValVallLysGluPheProSerTyrHisPheIleGlnHis 201  
QY 592 AAACGTGATATAACAAGCCACAACGAAGACTATACAAATCGTTGAGCAGTATGAA 645  
Db 202 ArgLeuGluLysThr--TyrValGluAspGlyGlyPheValGluGlnHisGlu 218  
RESULT 7  
US-09-977-897-3  
; Sequence 3, Application US/09977897  
; Patent No. 6780974  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Yih-Tai  
; APPLICANT: Cao, Longguang  
; TITLE OF INVENTION: A synthetic DNA encoding an orange seapen-derived green fluoresce  
; TITLE OF INVENTION: protein with codon preference of mammalian expression systems an  
; FILE REFERENCE: 41856-5  
; CURRENT APPLICATION NUMBER: US/09/977,897  
; CURRENT FILING DATE: 2001-10-15  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 224  
; TYPE: PRT  
; ORGANISM: Ptilosarcus gurneyi  
US-09-977-897-3  
Alignment Scores:  
Pred. No.: 2.08e-48 Length: 224  
Score: 473.50 Matches: 97  
Percent Similarity: 64.32% Conservative: 40  
Best Local Similarity: 45.54% Mismatches: 73  
Query Match: 38.97% Indels: 3  
DB: 4 Gaps: 3  
US-10-006-922A-11 (1-678) x US-09-977-897-3 (1-224)  
QY 13 AAGAATGTT---ATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGAACGGTCAAT 69  
Db 8 LysAsnThrGlyLeuLysGluIleMetSerAlaLysAlaSerValGluGlyIleValAsn 27  
QY 70 GGGCAGCAGTTTGAATAGAGCGGAAGGAGAGGGAGGCCCATACGAAGGCCACAAATACC 129  
Db 28 AsnHisValPheSerMetGluGlyPheGlyLysGlyAsnValLeuPheGlyAsnGlnLeu 47  
QY 130 GTAAAGCTTAAGTAAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATTTTGTCAACCA 189  
Db 48 MetGlnIleArgValThrLysGlyGlyProLeuProPheAlaPheAspIleValSerIle 67  
QY 190 CAATTTTCAGTATGAAGCAAGGTATATGTCAAGCACCTCGCCACATACAGACTATAAA 249  
Db 68 AlaPheGlnTyrGlyAsnArgThrPheThrLysTyrProAspAspIleAlaAspTyrPhe 87  
QY 250 AAGCTGTCAATTCCTGAAGGATTTAAATGGGAAGGGTTCATGAACCTTTGAAGACGGTGGC 309  
Db 88 ValGlnSerPheProAlaGlyPhePheTyrGluArgAsnLeuArgPheGluAspGlyAla 107  
QY 310 GTCGTTACTGTAAACCCAGGATTCAGTTTCAGGATGGCTGTTTCATCTACAAGGTCAAG 369  
Db 108 IleValAspIleArgSerAspIleSerLeuGluAspAspLysPheHisTyrLysValGlu 127  
QY 370 TTCATTGGCGTGAACCTTTCCTTCGATGGACCTGTTATGCAAAAGAACAAATGGGCTGG 429  
Db 128 TyrArgGlyAsnGlyPheProSerAsnGlyProValMetGlnLysAlaIleLeuGlyMet 147  
QY 430 GAAGCCAGCACTGAGCGTTTGTATCCTCGTATGGCGTGTGAAAGGAGAGATTTCATAAG 489  
Db 148 GluProSerPheGluValValTyrMetAsnSerGlyValLeuValGlyGluValAspLeu 167  
QY 490 GCTCTGAAGCTGAAGACGGTGGTCAATTACCTAGTTGAATTCAAAAGTATTACATGGCA 549  
Db 168 ValTyrLysLeuGluSerGlyAsnTyrTyrSerCysHisMetLysThrPheTyrArgSer 187  
QY 550 AAGAACCTGTG---CAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGATATAACA 606

Db 188 LysGlyGlyValLysGluPheProGluTyrHisPheIleHisArgLeuGluLysThr 207  
QY 607 AGCCACAACGAAGACTATACAAATCGTTGAGCAGTATGAA 645  
Db 208 --TyrValGluGluGlySerPheValGluGlnHisGlu 219  
RESULT 8  
US-09-977-897-13  
; Sequence 13, Application US/09977897  
; Patent No. 6780974  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Yih-Tai  
; APPLICANT: Cao, Longguang  
; TITLE OF INVENTION: A synthetic DNA encoding an orange seapen-derived green fluoresce  
; TITLE OF INVENTION: protein with codon preference of mammalian expression systems an  
; FILE REFERENCE: 41856-5  
; CURRENT APPLICATION NUMBER: US/09/977,897  
; CURRENT FILING DATE: 2001-10-15  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 13  
; LENGTH: 224  
; TYPE: PRT  
; ORGANISM: Ptilosarcus gurneyi  
US-09-977-897-13  
Alignment Scores:  
Pred. No.: 2.08e-48 Length: 224  
Score: 473.50 Matches: 97  
Percent Similarity: 64.32% Conservative: 40  
Best Local Similarity: 45.54% Mismatches: 73  
Query Match: 38.97% Indels: 3  
DB: 4 Gaps: 3  
US-10-006-922A-11 (1-678) x US-09-977-897-13 (1-224)  
QY 13 AAGAATGTT---ATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGAACGGTCAAT 69  
Db 8 LysAsnThrGlyLeuLysGluIleMetSerAlaLysAlaSerValGluGlyIleValAsn 27  
QY 70 GGGCAGCAGTTTGAATAGAGCGGAAGGAGAGGGAGGCCCATACGAAGGCCACAAATACC 129  
Db 28 AsnHisValPheSerMetGluGlyPheGlyLysGlyAsnValLeuPheGlyAsnGlnLeu 47  
QY 130 GTAAAGCTTAAGTAAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATTTTGTCAACCA 189  
Db 48 MetGlnIleArgValThrLysGlyGlyProLeuProPheAlaPheAspIleValSerIle 67  
QY 190 CAATTTTCAGTATGAAGCAAGGTATATGTCAAGCACCTCGCCACATACAGACTATAAA 249  
Db 68 AlaPheGlnTyrGlyAsnArgThrPheThrLysTyrProAspAspIleAlaAspTyrPhe 87  
QY 250 AAGCTGTCAATTCCTGAAGGATTTAAATGGGAAGGGTTCATGAACCTTTGAAGACGGTGGC 309  
Db 88 ValGlnSerPheProAlaGlyPhePheTyrGluArgAsnLeuArgPheGluAspGlyAla 107  
QY 310 GTCGTTACTGTAAACCCAGGATTCAGTTTCAGGATGGCTGTTTCATCTACAAGGTCAAG 369  
Db 108 IleValAspIleArgSerAspIleSerLeuGluAspAspLysPheHisTyrLysValGlu 127  
QY 370 TTCATTGGCGTGAACCTTTCCTTCGATGGACCTGTTATGCAAAAGAACAAATGGGCTGG 429  
Db 128 TyrArgGlyAsnGlyPheProSerAsnGlyProValMetGlnLysAlaIleLeuGlyMet 147  
QY 430 GAAGCCAGCACTGAGCGTTTGTATCCTCGTATGGCGTGTGAAAGGAGAGATTTCATAAG 489  
Db 148 GluProSerPheGluValValTyrMetAsnSerGlyValLeuValGlyGluValAspLeu 167  
QY 490 GCTCTGAAGCTGAAGACGGTGGTCAATTACCTAGTTGAATTCAAAAGTATTACATGGCA 549  
Db 168 ValTyrLysLeuGluSerGlyAsnTyrTyrSerCysHisMetLysThrPheTyrArgSer 187



```
QY 550 AAGAAGCCTGTG---CAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGATATAACA 606
    ||| ||| ::| ||| |||::: ||| ::|::: |||
Db 188 LysGlyGlyVallysGluPheProGluTyrHisPheIleHisArgLeuGluLysThr 207
    ::| |||::: ||| |||::: ||| |||::: |||
QY 607 AGCCACAACGAGACTATACAATCGTTGAGCAGTATGAA 645
    ::| |||::: ||| |||::: ||| |||::: |||
Db 208 ---TyrValGluGluGlySerPheValGluGlnHisGlu 219
    ::| |||::: ||| |||::: ||| |||::: |||

RESULT 9
US-09-977-897-14
; Sequence 14, Application US/09977897
; Patent No. 6780974
; GENERAL INFORMATION:
; APPLICANT: Chen, Yih-Tai
; APPLICANT: Cao, Longguang
; TITLE OF INVENTION: A synthetic DNA encoding an orange seapen-derived green fluoresce
; TITLE OF INVENTION: protein with codon preference of mammalian expression systems an
; FILE REFERENCE: 41856-5
; CURRENT APPLICATION NUMBER: US/09/977,897
; CURRENT FILING DATE: 2001-10-15
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Ptilosarcus gurneyi
US-09-977-897-14

Alignment Scores:
Pred. No.:      2.09e-48      Length:      225
Score:          473.50      Matches:      97
Percent Similarity: 64.32%      Conservative: 40
Best Local Similarity: 45.54%      Mismatches: 73
Query Match:    38.97%      Indels:      3
DB:              4      Gaps:      3

US-10-006-922A-11 (1-678) x US-09-977-897-14 (1-225)

QY 13 AAGAATGTT---ATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGGAACGGTCAAT 69
    ||||| |||::: ||||| ||| |||::: |||||
Db 8 LysAsnThrGlyLeuLysGluIleMetSerAlaLysAlaSerValGluGlyIleValAsn 27
    ::| |||::: ||| |||::: ||| |||::: |||
QY 70 GGGCACGAGTTTGAATAGAGGCGAAGGAGAGGGAGGCCATACGAAGGCCACAAATACC 129
    ||| ||| |||::: ||||| |||::: |||
Db 28 AsnHisValPheSerMetGluGlyPheGlyLysGlyAsnValLeuPheGlyAsnGlnLeu 47
    ::| |||::: ||| |||::: ||| |||::: |||
QY 130 GTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTGCTTGGGATATTTGTCACCA 189
    ::: |||::: ||| |||::: ||| |||::: |||
Db 48 MetGlnIleArgValThrLysGlyGlyProLeuProPheAlaPheAspIleValSerIle 67
    ::: |||::: ||| |||::: ||| |||::: |||
QY 190 CAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCAGACTATAAA 249
    ||||| |||::: ||| |||::: ||| |||::: |||
Db 68 AlaPheGlnTyrGlyAsnArgThrPheThrLysTyrProAspAspIleAlaAspTyrPhe 87
    ::| |||::: ||| |||::: ||| |||::: |||
QY 250 AAGCTGTCAATTCCTGAAGGATTTAAATGGGAAGGGTTCATGAACTTTGAAGACGGTGGC 309
    ||||| ||| ||| |||::: ||| |||::: |||
Db 88 ValGlnSerPheProAlaGlyPhePheTyrGluArgAsnLeuArgPheGluAspGlyAla 107
    ||||| ||| ||| |||::: ||| |||::: |||
QY 310 GTCGTTACTGTAAACCCAGGATTCAGATTGCGAGGATGGCTGTTTCATCTACAAGGTCAAG 369
    ::: ||| ||| |||::: ||| |||::: |||
Db 108 IleValAspIleArgSerAspIleSerLeuGluAspAspLysPheHisTyrLysValGlu 127
    ::| ||| ||| |||::: ||| |||::: |||
QY 370 TTCATTGGCGTGAACCTTCCTTCGATGGACCTGTATGCAAAAGAGCAATGGGCTGG 429
    ::| ||| ||| |||::: ||| |||::: |||
Db 128 TyrArgGlyAsnGlyPheProSerAsnGlyProValMetGlnLysAlaIleLeuGlyMet 147
    ::| ||| ||| |||::: ||| |||::: |||
QY 430 GAAGCCAGCACTAGCGGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAGATTTCATAAG 489
    ||| ||| ||| |||::: ||| |||::: |||
Db 148 GluProSerPheGluValValTyrMetAsnSerGlyValLeuValGlyGluValAspLeu 167
    ||| ||| ||| |||::: ||| |||::: |||
QY 490 GCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATTACATGGCA 549
    ||||| ||| ||| |||::: ||| |||::: |||
Db 168 valTyrLysLeuGluSerGlyAsnTyrTyrSerCysHisMetLysThrPheTyrArgSer 187
    ||||| ||| ||| |||::: ||| |||::: |||
```

```
QY 550 AAGAAGCCTGTG---CAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGATATAACA 606
    ||| ||| ::| ||| |||::: ||| ::|::: |||
Db 188 LysGlyGlyVallysGluPheProGluTyrHisPheIleHisArgLeuGluLysThr 207
    ::| |||::: ||| |||::: ||| |||::: |||
QY 607 AGCCACAACGAGACTATACAATCGTTGAGCAGTATGAA 645
    ::| |||::: ||| |||::: ||| |||::: |||
Db 208 ---TyrValGluGluGlySerPheValGluGlnHisGlu 219
    ::| |||::: ||| |||::: ||| |||::: |||

RESULT 10
US-09-977-897-15
; Sequence 15, Application US/09977897
; Patent No. 6780974
; GENERAL INFORMATION:
; APPLICANT: Chen, Yih-Tai
; APPLICANT: Cao, Longguang
; TITLE OF INVENTION: A synthetic DNA encoding an orange seapen-derived green fluorescer
; TITLE OF INVENTION: protein with codon preference of mammalian expression systems an
; FILE REFERENCE: 41856-5
; CURRENT APPLICATION NUMBER: US/09/977,897
; CURRENT FILING DATE: 2001-10-15
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Ptilosarcus gurneyi
US-09-977-897-15

Alignment Scores:
Pred. No.:      2.09e-48      Length:      226
Score:          473.50      Matches:      97
Percent Similarity: 64.32%      Conservative: 40
Best Local Similarity: 45.54%      Mismatches: 73
Query Match:    38.97%      Indels:      3
DB:              4      Gaps:      3

US-10-006-922A-11 (1-678) x US-09-977-897-15 (1-226)

QY 13 AAGAATGTT---ATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGGAACGGTCAAT 69
    ||||| |||::: ||||| ||| |||::: |||||
Db 8 LysAsnThrGlyLeuLysGluIleMetSerAlaLysAlaSerValGluGlyIleValAsn 27
    ::| |||::: ||| |||::: ||| |||::: |||
QY 70 GGGCACGAGTTTGAATAGAGGCGAAGGAGAGGGAGGCCATACGAAGGCCACAAATACC 129
    ||| ||| |||::: ||||| |||::: |||
Db 28 AsnHisValPheSerMetGluGlyPheGlyLysGlyAsnValLeuPheGlyAsnGlnLeu 47
    ::| |||::: ||| |||::: ||| |||::: |||
QY 130 GTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTGCTTGGGATATTTGTCACCA 189
    ::: |||::: ||| |||::: ||| |||::: |||
Db 48 MetGlnIleArgValThrLysGlyGlyProLeuProPheAlaPheAspIleValSerIle 67
    ::: |||::: ||| |||::: ||| |||::: |||
QY 190 CAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCAGACTATAAA 249
    ||||| |||::: ||| |||::: ||| |||::: |||
Db 68 AlaPheGlnTyrGlyAsnArgThrPheThrLysTyrProAspAspIleAlaAspTyrPhe 87
    ::| |||::: ||| |||::: ||| |||::: |||
QY 250 AAGCTGTCAATTCCTGAAGGATTTAAATGGGAAGGGTTCATGAACTTTGAAGACGGTGGC 309
    ||||| ||| ||| |||::: ||| |||::: |||
Db 88 valGlnSerPheProAlaGlyPhePheTyrGluArgAsnLeuArgPheGluAspGlyAla 107
    ||||| ||| ||| |||::: ||| |||::: |||
QY 310 GTCGTTACTGTAAACCCAGGATTCAGATTGCGAGGATGGCTGTTTCATCTACAAGGTCAAG 369
    ::: ||| ||| |||::: ||| |||::: |||
Db 108 IleValAspIleArgSerAspIleSerLeuGluAspAspLysPheHisTyrLysValGlu 127
    ::| ||| ||| |||::: ||| |||::: |||
QY 370 TTCATTGGCGTGAACCTTCCTTCGATGGACCTGTATGCAAAAGAGCAATGGGCTGG 429
    ::| ||| ||| |||::: ||| |||::: |||
Db 128 TyrArgGlyAsnGlyPheProSerAsnGlyProValMetGlnLysAlaIleLeuGlyMet 147
    ::| ||| ||| |||::: ||| |||::: |||
QY 430 GAAGCCAGCACTAGCGGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAGATTTCATAAG 489
    ||| ||| ||| |||::: ||| |||::: |||
Db 148 GluProSerPheGluValValTyrMetAsnSerGlyValLeuValGlyGluValAspLeu 167
    ||| ||| ||| |||::: ||| |||::: |||
QY 490 GCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATTACATGGCA 549
    ||||| ||| ||| |||::: ||| |||::: |||
Db 490 GCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATTACATGGCA 549
    ||||| ||| ||| |||::: ||| |||::: |||
```

```
Db 168 ValTyrLysLeuGluSerGlyAsnTyrTyrSerCysHisMetLysThrPheTyrArgSer 187
QY 550 AAGAAGCCTGTG---CAGCTACCAGGGTACTACTATGTTGACTCCAAACTGCATATAACA 606
Db 188 LysGlyGlyValLysGluPheProGluTyrHisPheIleHisHisArgLeuGluLysThr 207
QY 607 AGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAA 645
Db 208 ---TyrValGluGluGlySerPheValGluGlnHisGlu 219

RESULT 11
US-09-977-897-16
; Sequence 16, Application US/09977897
; Patent No. 6780974
; GENERAL INFORMATION:
; APPLICANT: Chen, Yih-Tai
; APPLICANT: Cao, Longguang
; TITLE OF INVENTION: A synthetic DNA encoding an orange seapen-derived green fluoresce
; TITLE OF INVENTION: protein with codon preference of mammalian expression systems an
; FILE REFERENCE: 41856-5
; CURRENT APPLICATION NUMBER: US/09/977,897
; CURRENT FILING DATE: 2001-10-15
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Ptilosarcus gurneyi
US-09-977-897-16

Alignment Scores:
Pred. No.: 2.1e-48 Length: 227
Score: 473.50 Matches: 97
Percent Similarity: 64.32% Conservative: 40
Best Local Similarity: 45.54% Mismatches: 73
Query Match: 38.97% Indels: 3
DB: 4 Gaps: 3

US-10-006-922A-11 (1-678) x US-09-977-897-16 (1-227)
QY 13 AAGAATGTT---ATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGCAACGGTCAAT 69
Db 8 LysAsnThrGlyLeuLysGluIleMetSerAlaLysAlaSerValGluGlyIleValAsn 27
QY 70 GGGCAGAGTTTGAATAAGGCGAAGGAGGGGAGGCCATACGAAGGCCACAATACC 129
Db 28 AsnHisValPheSerMetGluGlyPheGlyLysGlyAsnValLeuPheGlyAsnGlnLeu 47
QY 130 GTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATTTGTCACCA 189
Db 48 MetGlnIleArgValThrLysGlyGlyProLeuProPheAlaPheAspIleValSerIle 67
QY 190 CAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCAGACTATAAA 249
Db 68 AlaPheGlnTyrGlyAsnArgThrPheThrLysTyrProAspAspIleAlaAspTyrPhe 87
QY 250 AAGCTGTCAATTTCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAAGACGGTGGC 309
Db 88 ValGlnSerPheProAlaGlyPhePheTyrGluArgAsnLeuArgPheGluAspGlyAla 107
QY 310 GTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTACAAGGTCAAG 369
Db 108 IleValAspIleArgSerAspIleSerLeuGluAspAspLysPheHisTyrLysValGlu 127
QY 370 TTCATTGGCGTGAACCTTCCTCCGATGGACCTGTTATGCAAAAGAAGACATGGGCTGG 429
Db 128 TyrArgGlyAsnGlyPheProSerAsnGlyProValMetGlnLysAlaIleLeuGlyMet 147
QY 430 GAAGCCAGCACTGAGCGTTTGTATCCTCGTGTATGGCGTGTGAAAGGAGAGATTTCATAAG 489
Db 148 GluProSerPheGluValValTyrMetAsnSerGlyValLeuValGlyGluValAspLeu 167
QY 490 GCTCTGAAGCTGAAGACGGTGGTGCATTACCTAGTTGAATTCAAAAGTATTACATGGCA 549
```

```
Db 168 ValTyrLysLeuGluSerGlyAsnTyrTyrSerCysHisMetLysThrPheTyrArgSer 187
QY 550 AAGAAGCCTGTG---CAGCTACCAGGGTACTACTATGTTGACTCCAAACTGCATATAACA 606
Db 188 LysGlyGlyValLysGluPheProGluTyrHisPheIleHisHisArgLeuGluLysThr 207
QY 607 AGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAA 645
Db 208 ---TyrValGluGluGlySerPheValGluGlnHisGlu 219

RESULT 12
US-09-977-897-17
; Sequence 17, Application US/09977897
; Patent No. 6780974
; GENERAL INFORMATION:
; APPLICANT: Chen, Yih-Tai
; APPLICANT: Cao, Longguang
; TITLE OF INVENTION: A synthetic DNA encoding an orange seapen-derived green fluoresce
; TITLE OF INVENTION: protein with codon preference of mammalian expression systems an
; FILE REFERENCE: 41856-5
; CURRENT APPLICATION NUMBER: US/09/977,897
; CURRENT FILING DATE: 2001-10-15
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Ptilosarcus gurneyi
US-09-977-897-17

Alignment Scores:
Pred. No.: 2.1e-48 Length: 228
Score: 473.50 Matches: 97
Percent Similarity: 64.32% Conservative: 40
Best Local Similarity: 45.54% Mismatches: 73
Query Match: 38.97% Indels: 3
DB: 4 Gaps: 3

US-10-006-922A-11 (1-678) x US-09-977-897-17 (1-228)
QY 13 AAGAATGTT---ATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGCAACGGTCAAT 69
Db 8 LysAsnThrGlyLeuLysGluIleMetSerAlaLysAlaSerValGluGlyIleValAsn 27
QY 70 GGGCAGAGTTTGAATAAGGCGAAGGAGGGGAGGCCATACGAAGGCCACAATACC 129
Db 28 AsnHisValPheSerMetGluGlyPheGlyLysGlyAsnValLeuPheGlyAsnGlnLeu 47
QY 130 GTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATTTGTCACCA 189
Db 48 MetGlnIleArgValThrLysGlyGlyProLeuProPheAlaPheAspIleValSerIle 67
QY 190 CAATTTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCAGACTATAAA 249
Db 68 AlaPheGlnTyrGlyAsnArgThrPheThrLysTyrProAspAspIleAlaAspTyrPhe 87
QY 250 AAGCTGTCAATTTCTGAAGGATTTAAATGGGAAAGGGTCATGAACCTTTGAAGACGGTGGC 309
Db 88 ValGlnSerPheProAlaGlyPhePheTyrGluArgAsnLeuArgPheGluAspGlyAla 107
QY 310 GTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTACAAGGTCAAG 369
Db 108 IleValAspIleArgSerAspIleSerLeuGluAspAspLysPheHisTyrLysValGlu 127
QY 370 TTCATTGGCGTGAACCTTCCTCCGATGGACCTGTTATGCAAAAGAAGACATGGGCTGG 429
Db 128 TyrArgGlyAsnGlyPheProSerAsnGlyProValMetGlnLysAlaIleLeuGlyMet 147
QY 430 GAAGCCAGCACTGAGCGTTTGTATCCTCGTGTATGGCGTGTGAAAGGAGAGATTTCATAAG 489
Db 148 GluProSerPheGluValValTyrMetAsnSerGlyValLeuValGlyGluValAspLeu 167
```

```
QY 490 GCTCTGAAGCTGAAAGACGGTGGTCAATTACCTAGTTGAATTCAAAAGTATTATACATGGCA 549
    |||||::: ||| ::||| |||||: ||| :::
Db 168 ValTyrLysLeuGluSerGlyAsnTyrTyrSerCysHisMetLysThrPheTyrArgSer 187
    |||||::: ||| ::||| |||||: ||| :::
QY 550 AAGAAGCCTGTG---CAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGATATAACA 606
    |||||::: ||| ::||| |||||: ||| :::
Db 188 LysGlyGlyValLysGluPheProGluTyrHisPheIleHisHisArgLeuGluLysThr 207
    |||||::: ||| ::||| |||||: ||| :::
QY 607 AGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAA 645
    ::|||::: ||| ::||| |||||: ||| :::
Db 208 ---TyrValGluGluGlySerPheValGluGlnHisGlu 219
    ::|||::: ||| ::||| |||||: ||| :::
RESULT 13
US-09-977-897-18
; Sequence 18, Application US/09977897
; Patent No. 6780974
; GENERAL INFORMATION:
; APPLICANT: Chen, Yih-Tai
; APPLICANT: Cao, Longguang
; TITLE OF INVENTION: A synthetic DNA encoding an orange seapen-derived green fluoresce
; TITLE OF INVENTION: protein with codon preference of mammalian expression systems an
; FILE REFERENCE: 41856-5
; CURRENT APPLICATION NUMBER: US/09/977,897
; CURRENT FILING DATE: 2001-10-15
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Ptilosarcus gurneyi
US-09-977-897-18

Alignment Scores:
Pred. No.: 2.11e-48 Length: 229
Score: 473.50 Matches: 97
Percent Similarity: 64.32% Conservative: 40
Best Local Similarity: 45.54% Mismatches: 73
Query Match: 38.97% Indels: 3
DB: 4 Gaps: 3

US-10-006-922A-11 (1-678) x US-09-977-897-18 (1-229)
QY 13 AAGAATGTT---ATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGGAACGGTCAAT 69
    |||||::: ||| ::||| |||||: ||| :::
Db 8 LysAsnThrGlyLeuLysGluIleMetSerAlaLysAlaSerValGluGlyIleValAsn 27
    |||||::: ||| ::||| |||||: ||| :::
QY 70 GGGCAGAGTTTGAATAAGGCGAAGGAGGAGGGAGGCCATACGAAGGCCACAATACC 129
    |||||::: ||| ::||| |||||: ||| :::
Db 28 AsnHisValPheSerMetGluGlyPheGlyLysGlyAsnValLeuPheGlyAsnGlnLeu 47
    |||||::: ||| ::||| |||||: ||| :::
QY 130 GTAAAGCTTAAGTAAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATTTGTCACCA 189
    ::|||::: ||| ::||| |||||: ||| :::
Db 48 MetGlnIleArgValThrLysGlyGlyProLeuProPheAlaPheAspIleValSerIle 67
    ::|||::: ||| ::||| |||||: ||| :::
QY 190 CAATTTCAGTATGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCAGACTATAAA 249
    |||||::: ||| ::||| |||||: ||| :::
Db 68 AlaPheGlnTyrGlyAsnArgThrPheThrLysTyrProAspAspIleAlaAspTyrPhe 87
    |||||::: ||| ::||| |||||: ||| :::
QY 250 AAGCTGTCAATTCCTGAAGGATTTAAATGGGAAGGGTCATGAACCTTTGAAGACGGTGGC 309
    |||||::: ||| ::||| |||||: ||| :::
Db 88 ValGlnSerPheProAlaGlyPhePheTyrGluArgAsnLeuArgPheGluAspGlyAla 107
    |||||::: ||| ::||| |||||: ||| :::
QY 310 GTCGTTACTGTAAACCCAGGATTCAGATTTCAGGATGGCTGTTTCATCTACAAGGTCAAG 369
    ::|||::: ||| ::||| |||||: ||| :::
Db 108 IleValAspIleArgSerAspIleSerLeuGluAspAspLysPheHisTyrLysValGlu 127
    ::|||::: ||| ::||| |||||: ||| :::
QY 370 TTCATTGGCGTGAACCTTCCTCCGATGGACCTTTATGCAAAAGAGCAATGGGCTGG 429
    ::|||::: ||| ::||| |||||: ||| :::
Db 128 TyrArgGlyAsnGlyPheProSerAsnGlyProValMetGlnLysAlaIleLeuGlyMet 147
    ::|||::: ||| ::||| |||||: ||| :::
QY 430 GAAGCCAGCACTGAGCGTTTGTATCCTCGTGTGGCGTGTGAAAGGAGAGATTTCATAAG 489
    |||||::: ||| ::||| |||||: ||| :::
Db 148 GluProSerPheGluValValTyrMetAsnSerGlyValLeuValGlyGluValAspLeu 167
```

```
QY 490 GCTCTGAAGCTGAAAGACGGTGGTCAATTACCTAGTTGAATTCAAAAGTATTATACATGGCA 549
    |||||::: ||| ::||| |||||: ||| :::
Db 168 ValTyrLysLeuGluSerGlyAsnTyrTyrSerCysHisMetLysThrPheTyrArgSer 187
    |||||::: ||| ::||| |||||: ||| :::
QY 550 AAGAAGCCTGTG---CAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGATATAACA 606
    |||||::: ||| ::||| |||||: ||| :::
Db 188 LysGlyGlyValLysGluPheProGluTyrHisPheIleHisHisArgLeuGluLysThr 207
    |||||::: ||| ::||| |||||: ||| :::
QY 607 AGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAA 645
    ::|||::: ||| ::||| |||||: ||| :::
Db 208 ---TyrValGluGluGlySerPheValGluGlnHisGlu 219
    ::|||::: ||| ::||| |||||: ||| :::
RESULT 14
US-09-977-897-19
; Sequence 19, Application US/09977897
; Patent No. 6780974
; GENERAL INFORMATION:
; APPLICANT: Chen, Yih-Tai
; APPLICANT: Cao, Longguang
; TITLE OF INVENTION: A synthetic DNA encoding an orange seapen-derived green fluorescer
; TITLE OF INVENTION: protein with codon preference of mammalian expression systems an
; FILE REFERENCE: 41856-5
; CURRENT APPLICATION NUMBER: US/09/977,897
; CURRENT FILING DATE: 2001-10-15
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 230
; TYPE: PRT
; ORGANISM: Ptilosarcus gurneyi
US-09-977-897-19

Alignment Scores:
Pred. No.: 2.11e-48 Length: 230
Score: 473.50 Matches: 97
Percent Similarity: 64.32% Conservative: 40
Best Local Similarity: 45.54% Mismatches: 73
Query Match: 38.97% Indels: 3
DB: 4 Gaps: 3

US-10-006-922A-11 (1-678) x US-09-977-897-19 (1-230)
QY 13 AAGAATGTT---ATCAAGGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGGAACGGTCAAT 69
    |||||::: ||| ::||| |||||: ||| :::
Db 8 LysAsnThrGlyLeuLysGluIleMetSerAlaLysAlaSerValGluGlyIleValAsn 27
    |||||::: ||| ::||| |||||: ||| :::
QY 70 GGGCAGAGTTTGAATAAGGCGAAGGAGGAGGGAGGCCATACGAAGGCCACAATACC 129
    |||||::: ||| ::||| |||||: ||| :::
Db 28 AsnHisValPheSerMetGluGlyPheGlyLysGlyAsnValLeuPheGlyAsnGlnLeu 47
    |||||::: ||| ::||| |||||: ||| :::
QY 130 GTAAAGCTTAAGTAAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATTTGTCACCA 189
    ::|||::: ||| ::||| |||||: ||| :::
Db 48 MetGlnIleArgValThrLysGlyGlyProLeuProPheAlaPheAspIleValSerIle 67
    ::|||::: ||| ::||| |||||: ||| :::
QY 190 CAATTTCAGTATGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCAGACTATAAA 249
    |||||::: ||| ::||| |||||: ||| :::
Db 68 AlaPheGlnTyrGlyAsnArgThrPheThrLysTyrProAspAspIleAlaAspTyrPhe 87
    |||||::: ||| ::||| |||||: ||| :::
QY 250 AAGCTGTCAATTCCTGAAGGATTTAAATGGGAAGGGTCATGAACCTTTGAAGACGGTGGC 309
    |||||::: ||| ::||| |||||: ||| :::
Db 88 ValGlnSerPheProAlaGlyPhePheTyrGluArgAsnLeuArgPheGluAspGlyAla 107
    |||||::: ||| ::||| |||||: ||| :::
QY 310 GTCGTTACTGTAAACCCAGGATTCAGATTTCAGGATGGCTGTTTCATCTACAAGGTCAAG 369
    ::|||::: ||| ::||| |||||: ||| :::
Db 108 IleValAspIleArgSerAspIleSerLeuGluAspAspLysPheHisTyrLysValGlu 127
    ::|||::: ||| ::||| |||||: ||| :::
QY 370 TTCATTGGCGTGAACCTTCCTCCGATGGACCTTTATGCAAAAGAGCAATGGGCTGG 429
    ::|||::: ||| ::||| |||||: ||| :::
Db 128 TyrArgGlyAsnGlyPheProSerAsnGlyProValMetGlnLysAlaIleLeuGlyMet 147
    ::|||::: ||| ::||| |||||: ||| :::
QY 430 GAAGCCAGCACTGAGCGTTTGTATCCTCGTGTGGCGTGTGAAAGGAGAGATTTCATAAG 489
    |||||::: ||| ::||| |||||: ||| :::
```

Db 148 GluProSerPheGluValValTyrMetAsnSerGlyValLeuValGlyGluValAspLeu 167  
QY 490 GCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTGAATTCAAAAGTATTTACATGGCA 549  
Db 168 ValTyrLysLeuGluSerGlyAsnTyrTyrSerCysHisMetLysThrPheTyrArgSer 187  
QY 550 AAGAAGCCTGTG---CAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGATATAACA 606  
Db 188 LysGlyGlyValLysGluPheProGluTyrHisPheIleHisHisArgLeuGluLysThr 207  
QY 607 AGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAA 645  
Db 208 ---TyrValGluGluGlySerPheValGluGlnHisGlu 219

RESULT 15  
US-09-977-897-20  
; Sequence 20, Application US/09977897  
; Patent No. 6780974  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Yih-Tai  
; APPLICANT: Cao, Longguang  
; TITLE OF INVENTION: A synthetic DNA encoding an orange seapen-derived green fluoresce  
; TITLE OF INVENTION: protein with codon preference of mammalian expression systems an  
; FILE REFERENCE: 41856-5  
; CURRENT APPLICATION NUMBER: US/09/977,897  
; CURRENT FILING DATE: 2001-10-15  
; NUMBER OF SEQ.ID NOS: 27  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 20  
; LENGTH: 231  
; TYPE: PRT  
; ORGANISM: Ptilosarcus gurneyi  
US-09-977-897-20

Alignment Scores:  
Pred. No.: 2.12e-48 Length: 231  
Score: 473.50 Matches: 97  
Percent Similarity: 64.32% Conservative: 40  
Best Local Similarity: 45.54% Mismatches: 73  
Query Match: 38.97% Indels: 3  
DB: 4 Gaps: 3

US-10-006-922A-11 (1-678) x US-09-977-897-20 (1-231)  
QY 13 AAGAATGTT--ATCAAGGAGTTCATGAGGTTTAAGGTTTCGCATGGAAGGAACGGTCAAT 69  
Db 8 LysAsnThrGlyLeuLysGluIleMetSerAlaLysAlaSerValGluGlyIleValAsn 27  
QY 70 GGGCACGAGTTTGAATAGAAGGCGAAGGAGAGGGAGGCCATACGAAGGCCACAATACC 129  
Db 28 AsnHisValPheSerMetGluGlyPheGlyLysGlyAsnValLeuPheGlyAsnGlnLeu 47  
QY 130 GTAAAGCTTAAGGTAAACCAAGGGGGACCTTTGCCATTTCCTGGGATATTTGTCCACCA 189  
Db 48 MetGlnIleArgValThrLysGlyGlyProLeuProPheAlaPheAspIleValSerIle 67  
QY 190 CAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCAGACTATAAA 249  
Db 68 AlaPheGlnTyrGlyAsnArgThrPheThrLysTyrProAspAspIleAlaAspTyrPhe 87  
QY 250 AAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAAGACGGTGGC 309  
Db 88 ValGlnSerPheProAlaGlyPhePheTyrGluArgAsnLeuArgPheGluAspGlyAla 107  
QY 310 GTCGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTACAGGTCGAAG 369  
Db 108 IleValAspIleArgSerAspIleSerLeuGluAspLysPheHisTyrLysValGlu 127  
QY 370 TTCATTGGCGTGAACCTTTCCTCCGATGGACCTGTATATGCAAAAGAACACATGGGCTGG 429  
Db 128 TyrArgGlyAsnGlyPheProSerAsnGlyProValMetGlnLysAlaIleLeuGlyMet 147  
QY 430 GAAGCCAGCACTGAGCGCTTTGTATCCTCGTGTATGGCGGTGTTGAAAGGAGAGATTCTAAG 489

Db 148 GluProSerPheGluValValTyrMetAsnSerGlyValLeuValGlyGluValAspLeu 167  
QY 490 GCTCTGAAGCTGAAAGACGGTGGTCATTACCTAGTGAATTCAAAAGTATTTACATGGCA 549  
Db 168 ValTyrLysLeuGluSerGlyAsnTyrTyrSerCysHisMetLysThrPheTyrArgSer 187  
QY 550 AAGAAGCCTGTG---CAGCTACCAGGGTACTACTATGTTGACTCCAAACTGGATATAACA 606  
Db 188 LysGlyGlyValLysGluPheProGluTyrHisPheIleHisHisArgLeuGluLysThr 207  
QY 607 AGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAA 645  
Db 208 ---TyrValGluGluGlySerPheValGluGlnHisGlu 219

Search completed: June 30, 2005, 16:06:40  
Job time : 36 secs



**THIS PAGE BLANK (USPIU)**